PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-362559
(43)Date of publication of application : 24.12.2004

(51)Int.Cl. G06F 13/00

(21)Application number : 2004-148159
(22)Date of filing : 18.05.2004

(71)Applicant : MICROSOFT CORP
(72)Inventor : GOODMAN JOSHUA T
ROUNTHWAITE ROBERT L
GWOZDZ DANIEL
MEHR JOHN D
HOWELL NATHAN D
RUPERSBURG MICAH C
STARBUCK BRYAN T

(30)Priority
Priority number : 2003 454168 Priority date : 04.06.2003 Priority country : US

(54) FEATURES AND LIST OF ORIGINATION AND DESTINATION FOR SPAM PREVENTION

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a system and method that facilitates extracting data from messages for spam filtering.
SOLUTION: The extracted data can be treated in the form of features, and the features can be used with machine learning system so that improved filters can be built. Origination information and such data as to be related to other information embedded in message body which enables a receipient of messages to get in touch and/or response to the sender of the messages can be extracted as features. The features or a subset can be deobfuscated and/or normalized, before they are used as features of the machine learning system. The features can be used to populate several feature lists which

facilitates prevention and detection of spams. Examplary features includes an e-mail address, an IP addresses, a URL, embedded images indicating a URL and/or parts of them.

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号
特開2004-362559
(P2004-362559A)

(43)公開日 平成16年12月24日(2004.12.24)

(51)Int.Cl.[7]	F I	テーマコード（参考）
G06F 13/00	G06F 13/00 610Q	

審査請求 未請求 請求項の数 54 ＯＬ 外国語出願 （全 38 頁）

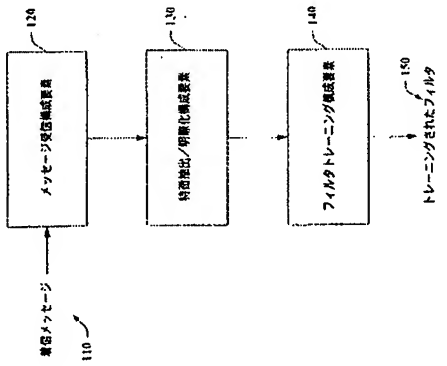
(21)出願番号 特願2004-148159(P2004-148159)
(22)出願日 平成16年5月18日(2004.5.18)
(31)優先権主張番号 10/454,168
(32)優先日 平成15年6月4日(2003.6.4)
(33)優先権主張国 米国(US)

(71)出願人 500046438
マイクロソフト コーポレーション
アメリカ合衆国 ワシントン州 98052-6399 レッドモンド ワン マイクロソフト ウェイ
(74)代理人 100077481
弁理士 谷 義一
(74)代理人 100088915
弁理士 阿部 和夫
(72)発明者 ジョシュア ティー. グッドマン
アメリカ合衆国 98052 ワシントン州 レッドモンド ノースイースト 38 ストリート 17424

最終頁に続く

(54)【発明の名称】スパム防止のための送信元および宛先の特徴およびリスト

(57)【要約】
【課題】 スパムフィルタリングのためにメッセージからデータを抽出することを容易にするシステムおよび方法を提供すること。
【解決手段】 抽出されたデータは特徴の形で扱うことができ、特徴は、機械学習システムと共に使用し、改良されたフィルタを構築することができる。送信元情報、ならびにメッセージの受信者がメッセージの送信者に接触および／または応答することを可能にするメッセージの本文内に埋め込まれた他の情報に関連付けられたデータを特徴として抽出することができる。特徴またはそのサブセットは、機械学習システムの特徴として使用する前に正規化および／または明瞭化することができる。特徴は、スパム検出および防止を容易にする複数の特徴リストに配置するために使用することができる。例示的な特徴には、電子メールアドレス、ＩＰアドレス、ＵＲＬ、ＵＲＬを指す埋込み画像、および／またはその一部分が含まれる。
【選択図】 図1



【特許請求の範囲】
【請求項1】
スパム処理に関連するデータの抽出を容易にするシステムであって、
アイテムを受信し、メッセージの送信元もしくは該送信元の一部分、および／または意図される受信者が該メッセージに関連して接触、応答、または受信することを可能にする情報に関連付けられた特徴のセットを抽出する構成要素と、
フィルタの構築に関連して、前記抽出された特徴のサブセットを使用する構成要素と
を備えたことを特徴とするシステム。
【請求項2】
前記特徴のサブセットを明瞭化する正規化構成要素をさらに備えたことを特徴とする請求項１に記載のシステム。
【請求項3】
前記フィルタは、スパムフィルタであることを特徴とする請求項１に記載のシステム。
【請求項4】
前記フィルタは、ペアレンタル制御フィルタであることを特徴とする請求項１に記載のシステム。
【請求項5】
前記明瞭化された特徴を使用して、スパムおよび非スパムのうち少なくとも１つを学習する機械学習システム構成要素をさらに備えたことを特徴とする請求項１に記載のシステム。
【請求項6】
前記特徴のサブセットは、少なくとも１つのＩＰアドレスを含み、前記少なくとも１つのＩＰアドレスは、メッセージ内に位置する返信先アドレス、カーボンコピーアドレス、ｍａｉｌｔｏ：アドレス、発信元アドレス、およびＵＲＬのいずれか１つの少なくとも一部分であることを特徴とする請求項１に記載のシステム。
【請求項7】
前記ＩＰアドレスは、ブロックＩＤを含み、前記ブロックＩＤは、少なくとも１つの特徴として抽出することができることを特徴とする請求項６に記載のシステム。
【請求項8】
前記ブロックＩＤは、少なくとも部分的には、ブロックディレクトリを調べることによって決定されることを特徴とする請求項７に記載のシステム。
【請求項9】
前記ブロックディレクトリは、ａｒｉｎ．ｎｅｔであることを特徴とする請求項８に記載のシステム。
【請求項10】
前記ブロックＩＤは、少なくとも部分的には、推測することによって決定され、該決定により、前記ＩＰアドレスの少なくとも最初の１ビット、少なくとも最初の２ビット、少なくとも最初の３ビット、および少なくとも最大で最初の３１ビットまでのいずれか１つを特徴として抽出することを特徴とする請求項７に記載のシステム。
【請求項11】
前記特徴のサブセットは、ＩＰアドレスの最初の１ビットから最初の３１ビットまでの各々を含むことを特徴とする請求項１に記載のシステム。
【請求項12】
前記特徴のサブセットは、ＵＲＬを含むことを特徴とする請求項１に記載のシステム。
【請求項13】
前記ＵＲＬアドレスは、前記メッセージの本文内に配置されたもの、前記メッセージ内でテキストとして埋め込まれているもの、前記メッセージ内で画像として埋め込まれているもののうち少なくとも１つであることを特徴とする請求項１２に記載のシステム。
【請求項14】
前記抽出された特徴の少なくともサブセットを使用して少なくとも１つの特徴リストに

配置する構成要素をさらに備えたことを特徴とする請求項１に記載のシステム。
【請求項15】
前記少なくとも１つの特徴リストは、問題のないユーザのリスト、スパム送信者のリスト、本物の送信者を示す肯定的な特徴のリスト、スパムを示す特徴のリストのいずれか１つであることを特徴とする請求項１４に記載のシステム。
【請求項16】
前記特徴のサブセットは、少なくとも１つのＵＲＬを含むことを特徴とする請求項１に記載のシステム。
【請求項17】
前記ＵＲＬは、前記メッセージの本文内にテキストとして埋め込まれていることを特徴とする請求項１６に記載のシステム。
【請求項18】
前記ＵＲＬは、前記メッセージの本文内のリンクの少なくとも一部分であることを特徴とする請求項１６に記載のシステム。
【請求項19】
前記ＵＲＬは、メッセージ内の画像として埋め込まれたリンクの少なくとも一部分であることを特徴とする請求項１６に記載のシステム。
【請求項20】
前記特徴のサブセットは、電子メールアドレスから抽出されたホスト名およびドメイン名のうち少なくとも１つを含むことを特徴とする請求項１に記載のシステム。
【請求項21】
前記特徴のサブセットは、電子メールアドレスおよびＵＲＬのいずれか１つから抽出されたＦＱＤＮの少なくとも一部分を含むことを特徴とする請求項１に記載のシステム。
【請求項22】
前記特徴のサブセットは、電子メールアドレスおよびＵＲＬのいずれか１つから抽出されたドメイン名の少なくとも一部分を含むことを特徴とする請求項１に記載のシステム。
【請求項23】
前記抽出された前記特徴のサブセットの少なくとも一部分は、機械学習システムと共に使用される前に正規化されることを特徴とする請求項１に記載のシステム。
【請求項24】
前記抽出された前記特徴のサブセットの少なくとも一部分は、少なくとも１つの特徴リストに配置するために使用される前に正規化されることを特徴とする請求項１に記載のシステム。
【請求項25】
ＵＲＬ、電子メールアドレス、ＩＰアドレスのうち少なくとも１つの少なくとも一部分を、成人、成人向け内容、ふさわしくない、一部の年齢にとってふさわしくない、あらゆる年齢にとってふさわしい、不適切、および適切のうちのいずれか１つとして分類する分類構成要素をさらに備えたことを特徴とする請求項１に記載のシステム。
【請求項26】
前記分類構成要素は、ペアレンタル制御システムであることを特徴とする請求項２５に記載のシステム。
【請求項27】
前記分類構成要素は、少なくとも１つの特徴の型を、前記ＵＲＬ、ウェブサイトアドレス、前記ＩＰアドレスのうち少なくとも１つの前記分類された一部分に割り当てることを特徴とする請求項２５に記載のシステム。
【請求項28】
前記特徴のセットは、少なくとも１つの非無料電話番号を含み、前記電話番号は、前記メッセージに関連付けられた送信者または連絡先の地理的位置をマッピングすることを容易にするエリアコードを含むことを特徴とする請求項１に記載のシステム。
【請求項29】

請求項１のコンピュータ実行可能構成要素を格納することを特徴とするコンピュータ読取可能な媒体。
【請求項30】
請求項１に記載の前記システムを使用することを特徴とするコンピュータ。
【請求項31】
スパム処理に関連するデータの抽出を容易にする方法であって、
メッセージを受信するステップと、
前記メッセージの送信元もしくはその一部、および／または意図される受信者が前記メッセージに関連して接触、応答、または受信することを可能にする情報に関連付けられた特徴のセットを抽出するステップと、
フィルタを構築することに関連して、前記抽出された特徴のサブセットを使用するステップと
を備えたことを特徴とする方法。
【請求項32】
前記特徴のセットは、ＩＰアドレスの少なくとも一部分を含むことを特徴とする請求項３１に記載の方法。
【請求項33】
前記ＩＰアドレスの少なくとも一部分を抽出するステップは、
ブロックＩＤが追加の特徴として抽出されるように、ブロックＩＤディレクトリを調べ、前記ＩＰアドレスに対応する少なくとも１つのブロックＩＤを決定する動作、および
前記ＩＰアドレスから少なくとも最初の１ビットから最初の３１ビットまでの各々を抽出する動作のうち少なくとも１つを実行するステップを含むことを特徴とする請求項３２に記載の方法。
【請求項34】
少なくとも１つの抽出されたＩＰアドレスは、少なくとも１つのサーバに対応することを特徴とする請求項３２に記載の方法。
【請求項35】
前記少なくとも１つのサーバを追加の特徴として抽出するステップをさらに備えたことを特徴とする請求項３４に記載の方法。
【請求項36】
前記メッセージから抽出された特徴の少なくともサブセットを明瞭化するステップをさらに備えたことを特徴とする請求項３１に記載の方法。
【請求項37】
前記メッセージから抽出された少なくとも１つの特徴の少なくとも一部分を明瞭化するステップをさらに備えたことを特徴とする請求項３１に記載の方法。
【請求項38】
前記メッセージから抽出された発信元ＩＰアドレスを明瞭化するステップは、添付先（ａｐｐｅｎｄｅｄ－ｔｏ）ＩＰアドレスの同一性を検証するため、複数の添付先ＩＰアドレスを遡って追跡するステップを含むことを特徴とする請求項３７に記載の方法。
【請求項39】
１度に少なくとも１つの接尾語を除去し、該除去により、それぞれの追加の特徴を生じる動作と、
１度に少なくとも１つの接頭語を除去し、該除去により、それぞれの追加の特徴を生じる動作とのうち、少なくとも１つを実行するステップを含む、ウェブサイトアドレスから追加の特徴を抽出するステップをさらに備えたことを特徴とする請求項３７に記載の方法。
【請求項40】
前記特徴のセットは、返信先アドレス、カーボンコピーアドレス、ｍａｉｌｔｏ：アドレス、ＵＲＬ、リンク、および発信元アドレスのいずれか１つの少なくとも一部分を含むことを特徴とする請求項３７に記載の方法。

【請求項41】
前記抽出された特徴の少なくともサブセットは、前記メッセージの本文内でテキストおよび画像のうち１つとして埋め込まれていることを特徴とする請求項３１に記載の方法。
【請求項42】
前記特徴のセットは、ホスト名およびドメイン名を含むことを特徴とする請求項３１に記載の方法。
【請求項43】
前記メッセージに関連付けられたふさわしい内容、およびふさわしくない内容のいずれか１つを示すため、１つまたは複数の抽出された特徴および／またはその一部分を分類するステップと、該分類を追加の特徴として使用するステップとをさらに備えたことを特徴とする請求項３１に記載の方法。
【請求項44】
少なくとも部分的にはそれぞれの抽出された特徴に基づいて、メッセージ内容をユーザに通知するため、それぞれの抽出された特徴に特徴の型を割り当てるステップと、前記特徴の型を追加の特徴として使用するステップとをさらに備えたことを特徴とする請求項３１に記載の方法。
【請求項45】
特徴の型および特徴の少なくとも１つが、希少性および共通性のいずれか１つであることを決定するステップと、特徴の希少性および共通性を追加の特徴として使用するステップとをさらに備えたことを特徴とする請求項４４に記載の方法。
【請求項46】
前記特徴のサブセットは、機械学習システムを介してフィルタを構築することに関連して使用されることを特徴とする請求項３１に記載の方法。
【請求項47】
前記フィルタは、スパムフィルタであることを特徴とする請求項３１に記載の方法。
【請求項48】
前記フィルタは、ペアレンタル制御フィルタであることを特徴とする請求項３１に記載の方法。
【請求項49】
前記メッセージから抽出された特徴の少なくともサブセットを使用して１つまたは複数の特徴リストに配置するステップをさらに含むことを特徴とする請求項３１に記載の方法。
【請求項50】
特徴リストは、非スパム送信者を示す肯定的な特徴リスト、およびスパム送信者を示す悪質な特徴リストのうち少なくとも１つ含むことを特徴とする請求項４９に記載の方法。
【請求項51】
前記抽出された特徴は、少なくとも部分的には機械学習システムの特徴として使用される前に、明瞭化されることを特徴とする請求項３１に記載の方法。
【請求項52】
前記抽出された特徴は、少なくとも部分的には特徴リストに配置する特徴として使用される前に、明瞭化されることを特徴とする請求項３１に記載の方法。
【請求項53】
メッセージからデータを抽出することを容易にして、複数のコンピュータプロセス間で送信されるように適合されたデータパケットであって、
メッセージを受信すること、前記メッセージの送信元もしくは該送信元の一部、および／または意図される受信者が前記メッセージに関連して接触、応答、または受信することを可能にする情報に関連付けられた特徴のセットを抽出することと、フィルタの構築に関連して前記抽出された特徴のサブセットを使用することとに関連付けられた情報を
備えたことを特徴とするデータパケット。
【請求項54】

スパム処理に関連するデータを抽出することを容易にするシステムであって、
メッセージを受信するための手段と、
前記メッセージの送信元もしくは該送信元の一部、および／または意図される受信者が前記メッセージに関連して接触、応答、または受信することを可能にする情報に関連付けられた特徴のセットを抽出するための手段と、
フィルタの構築に関連して、前記抽出された特徴のサブセットを使用する手段と
を備えたことを特徴とするシステム。

【発明の詳細な説明】
【技術分野】
【0001】
本発明は、本物のメール（すなわち、問題のないメール）および望ましくないメールの双方を識別するシステムおよび方法に関し、より詳細には、スパム防止を容易にするために、電子メッセージを処理してデータを抽出するためのシステムおよび方法に関する。
【背景技術】
【0002】
インターネットなど地球全体の通信ネットワークの登場により、膨大な数の潜在する顧客に情報を届ける商業的な機会が提供された。電子メッセージング、特に電子メール（ｅメール）は、（「スパム」とも呼ばれる）望まれてもいない広告および勧誘をネットワークユーザにまき散らすための手段として、ますます広まりつつある。
【0003】
コンサルティングおよび市場調査会社であるＲａｄｉｃａｔｉ　Ｇｒｏｕｐ，Ｉｎｃ．は、２００２年８月現在、毎日２０億通のジャンク電子メールメッセージが送られており、この数は２年毎に３倍になると予想されると見積もっている。個人および団体（たとえば、企業、政府機関）はますます、ジャンクメッセージによって迷惑を受け、しばしば不快な思いをさせられている。したがって、スパムは現在、またはまもなく、信頼できるコンピューティングを揺るがす大きな脅威になるであろう。
【0004】
スパムを防ぐために使用される主な技法は、フィルタリングシステム／方法を使用することである。実証済みのフィルタリング技法の１つは、機械学習手法に基づくものであり、機械学習フィルタは、そのメッセージがスパムである確率を着信メッセージに割り当てる。この手法では、一般に２種類の事例メッセージ（たとえば、スパムメッセージと非スパムメッセージ）から特徴（ｆｅａｔｕｒｅ）が抽出され、この２つの種類間を確率的に弁別するために学習フィルタが適用される。多数のメッセージの特徴は内容（たとえば、メッセージの主題および／または本文内の単語または句）に関係するため、そのような型のフィルタは、一般に「コンテンツベースのフィルタ」と呼ばれる。
【発明の開示】
【発明が解決しようとする課題】
【0005】
そのようなスパムフィルタリング技法の猛反撃を受け、多数のスパム送信者は、スパムフィルタを回避および／または迂回するためにその正体を隠す方法を考えるようになっている。したがって、従来のコンテンツベースのフィルタや適応型フィルタは、偽装されたスパムメッセージを認識し遮断する上で効果を失うおそれがある。
【課題を解決するための手段】
【0006】
本発明のいくつかの態様についての基本的な理解を得るため、以下に本発明を簡単にまとめる。この概要は、本発明の広範な全体像を示すものではない。本発明の主な／決定的な要素を特定すること、あるいは本発明の範囲を説明することは意図されていない。後から提供されるより詳しい説明の序文として、本発明のいくつかの概念を簡単な形態で示すことを目的とするにすぎない。
【0007】

スパム送信者は、そのメッセージ内のほとんどすべての情報を偽装することができる。たとえば、機械学習システムにとって特徴として使用される単語そのものが存在しないように、画像を埋め込むことができる。画像を所定の方法で歪ませて、ＯＣＲソフトウェアを使用することが困難になるか、または少なくとも時間がかかるようにすることもできる。しかし、どれだけ多数の特徴を除去しても、依然として有用な情報はある。第１に、スパム送信者は、どこからかメッセージを送らなければならない。どのＩＰアドレスからメッセージが届けられたか検出することができるのである。第２に、スパム送信者は大抵何かを販売しようとしており、したがって接触する方法が含まれているはずである。これは無料電話番号である可能性もあるが、スパム送信者は、苦情からコストが高くつくため、これを使用するのを避けるかもしれない。無料でない電話番号である可能性もあるが、逆に応答率がより低いため、そうしないかもしれない。またはこれに替えて、ＵＲＬである可能性がある（たとえば、http://www.spamcorp.com/buyenlarger.htm）。このＵＲＬが、フィルタおよび／またはソフトウェアによって検出されるのをより困難にするために、画像内に埋め込まれている可能性がある。しかし、スパム送信者は、ユーザがＵＲＬをブラウザにタイプすることを必要とし、それにより応答率が低下する可能性があるため、そうしたがらないであろう。

【0008】

スパム送信者が接触を受ける最も可能性の高い方法は、埋込みリンク、または何らかの埋込み電子メールアドレスによるものである。たとえば、「もっと知るにはここをクリックしてください」であり、「ここをクリック」には、本発明の一態様に従って機械学習システムが検出および使用することができる特定のウェブページへのリンクが含まれている。同様に、そこに返信すべきアドレス（たとえば、典型的には「発信元アドレス（ｆｒｏｍ　ａｄｄｒｅｓｓ）」であるが、「返信先（ｒｅｐｌｙ－ｔｏ）」アドレスのある場合がある）、または任意の埋込みｍａｉｌｔｏ：リンク（リンク上でクリックすることによってメールメッセージを送ることを可能にするリンク）、若しくは任意の他の埋込み電子メールアドレスである。さらに、スパム送信者は、しばしばメッセージ内に画像を含める。大きな画像を何回もメールするのはコストがかかるため、スパム送信者は、しばしばその画像への特別なリンクだけ埋め込み、そのリンクによってその画像のダウンロードが行われてしまう。これらのリンクが指す位置もまた、特徴として使用することができる。

【0009】

メール発信元アドレス、メール返信先アドレス、埋込みｍａｉｌｔｏ：アドレス、外部リンク、および外部画像のリンクから引き出された情報に関して、そのような情報の少なくとも一部分を機械学習システムの特徴として使用することができ、重みまたは確率が関連付けられ、あるいは情報をリストに追加することができる。たとえば、スパムだけ、または問題のないメールだけ、または９０％を超える問題のないメールなどを送信するＩＰアドレスまたは発信元アドレスのリストを保つことができる。特定のリンクまたはアドレスがそのようなリスト上にあるという事実を、機械学習システムの特徴として、または任意の他のスパムフィルタリングシステムの一部として、あるいはその両方として使用することができる。

【0010】

本発明は、メッセージの特定の一部分を調べることにより、偽装されたスパムメッセージを識別することを容易にするシステムおよび方法を提供する。より具体的には、本発明は、本物のメッセージからスパムメッセージを区別するために、電子メール（ｅメール）などメッセージを処理して送信元および／または宛先データを抽出するものである。この処理は、ＩＰアドレス情報、電子メールアドレス情報、および／またはユニフォームリソースロケータ（ＵＲＬ）情報を識別して構文解析するための、および抽出されたデータをスパム属性（例えば、問題のないユーザと悪質なユーザ、または問題のない送信者と悪質な送信者）に関連付けるための様々な技法を含む。悪質なユーザまたは悪質な送信者は、たとえばスパム送信者（すなわち、スパムを送る者）と見なされるはずである。

【0011】

抽出されたデータ、または少なくともその一部分を使用し、機械学習システム用の特徴セットを生成することができる。機械学習技法は、メッセージの内容を調べ、そのメッセージがスパムであるかどうか判定する。スパム送信者は、処理するのが困難な画像内にその情報のほとんどを入れるなどのことをすることによって、メッセージの内容のほとんどを不明瞭にすることができる。しかし、スパム送信者は、受信者がスパム送信者に容易に接触する何らかの方法を提供することが必要なため、メッセージの送信源を完全に偽装することができない。そのようなものとしては、リンク（たとえば、ＵＲＬ）および／または電子メールアドレス（たとえば、ＩＰアドレス）の使用などがある。これらの型の情報または変形形態、あるいはその一部分を、スパム検出器の特徴として使用することができる。具体的には、たとえば機械学習システムにより、その情報を使用してスパム検出器および／またはスパムフィルタをトレーニングすることができる。

【0012】

本発明はまた、ペアレンタル制御システムと共に用いることができる。ペアレンタル制御システムは、メッセージが不適切であることをユーザに通知することができ、および「ポルノ題材を含む」など、その理由を示すこともできる。本発明の一態様によれば、１つまたは複数の抽出され正規化された特徴（たとえば、ＵＲＬ）をペアレンタル制御システムまたはフィルタに通し、ペアレンタル制御システムの分類を得ることができる。この分類を機械学習システムの追加の特徴として使用し、スパムフィルタを構築および／または改善することを容易にすることができる。

【0013】

さらに、抽出された特徴を種類によって分類し、スパム度に従って加重し、および肯定的（たとえば、スパムでない可能性がより高い）特徴または否定的（たとえば、スパムである可能性がより高い）特徴と指定することができる。また、特徴群を使用し、たとえば、非スパム送信者リストおよびスパム送信者リストなどのリストを作成することができる。

【0014】

前述の、および関連する目的を達成するため、本明細書では、本発明のある種の例示的な実施形態が、以下の説明および添付の図面と共に説明される。しかし、これらの実施形態は、本発明の原理を使用することができる様々な方法のいくつかを示すものにすぎず、本発明は、そのような実施形態とその等価なすべてを含むものとする。本発明の他の利点および新規の特徴は、図面と共に考察し、以下の本発明の詳細な説明から明らかにすることができる。

【発明を実施するための最良の形態】

【0015】

次に、本発明について図面を参照しながら説明する。図面では、全体を通して同じ要素を参照するために同じ符号が使用される。以下の説明では、あくまで説明の目的で、本発明を十分理解するための多数の具体的な詳細が説明される。しかし、これら具体的な詳細なしに本発明を実施することができることは自明であることが理解される。場合によっては、本発明の説明を容易にするため、周知の構造およびデバイスがブロック図の形態で示されている。

【0016】

本願では、「構成要素」および「システム」という用語は、ハードウェア、ハードウェアとソフトウェアの組合せ、ソフトウェア、または実行中のソフトウェアのいずれかであるコンピュータ関連のエンティティを指すものとする。たとえば、構成要素は、それだけには限らないが、プロセッサ上で動作するプロセス、プロセッサ、オブジェクト、実行可能物、実行のスレッド、プログラム、および／またはコンピュータとすることができる。例示のため、サーバ上で動作するアプリケーションもサーバも共に構成要素とすることができる。プロセスおよび／または実行のスレッド内に１つまたは複数の構成要素が常駐することができ、構成要素は、１つのコンピュータに局所化する、および／または複数のコンピュータ間で分散することができる。

【0017】
本発明は、機械学習式スパムフィルタリング用のトレーニングデータを生成することに関連して、様々な推理スキームおよび／または技法を組み込むことができる。本明細書では、「推理」という用語は、概して、事象および／またはデータを介して取り込まれた観察結果の集合から、システムの状態、環境、および／またはユーザについて推論する工程、またはこれらの状態を推理する工程を指す。推理は、たとえば、特定の内容または動作を識別するために使用することができ、あるいは状態全体にわたって確率分布を生成することができる。推理は、確率的なもの、すなわち、データおよび事象の考察に基づいた当該の状態全体にわたる確率分布の計算とすることができる。推理はまた、事象および／またはデータの集合から、より高いレベルの事象を構成するために使用される技法を指す可能性がある。そのような推理により、事象群が時間的に近接しているか否か、およびその事象およびデータが１つまたは複数の事象源やデータ源からのものであるか否かにかかわらず、観察された事象および／または記憶された事象データの集合から新しい事象または動作が構築される。
【0018】
本明細書全体にわたって「メッセージ」という用語が広く使用されているが、そのような用語は、電子メールそれ自体に限定されず、任意の好適な通信アーキテクチャを介して配布することができるどの形態の電子メッセージングをも含むように適切になすことができることを理解されたい。たとえば、２名以上での会議を容易にする会議アプリケーション（たとえば、対話型チャットプログラムおよびインスタントメッセージングプログラム）もまた、本明細書に開示されているフィルタリングの利点を活用することができる。というのは、望ましくないテキストは、ユーザがメッセージを交換しているとき、通常のチャットメッセージ内に電子的にばらまかれ、および／またはリードオフメッセージ、クロージングメッセージ、もしくは上記のすべてとして挿入される可能性があるからである。この特定のアプリケーションでは、望ましくない内容（たとえば、コマーシャル、勧誘、または広告）を取り込み、スパムとしてタグ付けするために、特定のメッセージ内容（テキストおよび画像）を自動的にフィルタリングするようにフィルタをトレーニングすることができるのである。
【0019】
本発明では、「受信者」という用語は、着信メッセージまたはメール項目の受取人を指す。「ユーザ」という用語は、状況に応じて、受信者または送信者を指す可能性がある。たとえば、ユーザは、状況や用語の適用に応じて、スパムを送信する電子メールユーザを指す可能性があり、および／またはスパムを受信する電子メール受信者を指す可能性がある。
【0020】
インターネットプロトコル（ＩＰ）アドレスは、一般にインターネット上の装置を表す３２ビットの数値である。これらの数値は、２台の装置の間で通信するとき使用される。これらの数値は、一般に「ｘｘｘ．ｘｘｘ．ｘｘｘ．ｘｘｘ」の形態で表され、各ｘｘｘは、０と２５５の間である。残念ながら、ＩＰアドレスは覚えておくことが困難である。そのため、「ドメイン名」および「ホスト名」という取決めが作成されている。「ドメイン名」は、インターネット上の装置のグループ（おそらくは、単一の装置）の名前であり、一般に「ｘ．ｃｏｍ」または「ｙ．ｅｄｕ」または「ｃｏｕｒｔｓ．ｗａ．ｇｏｖ」という形態のものである。
【0021】
完全修飾ドメイン名（ＦＱＤＮ）は、インターネット上の特定の装置、たとえば「ｂ．ｘ．ｃｏｍ」または「ｃ．ｙ．ｅｄｕ」または「ｗｗｗ．ｃｏｕｒｔｓ．ｗａ．ｇｏｖ」であり、ドメイン名部分は、それぞれ「ｘ．ｃｏｍ」または「ｙ．ｅｄｕ」または「ｃｏｕｒｔｓ．ｗａ．ｇｏｖ」である。「ｂ」「ｃ」「ｗｗｗ」部分はそれぞれ、ＦＱＤＮのホスト名部分と呼ばれる。一般に、ＩＰアドレスは、ドメイン名を使用することができるどの状況でも使用することができる（たとえば、ＤＮ／ＩＰは、どちらの可能性もあるこ

とを示す）。また、一般に、ＩＰアドレスは、ＦＱＤＮを使用することができるどの状況でも使用することができる（たとえば、ＦＱＤＮ／ＩＰは、どちらの可能性もあることを示す）。電子メールアドレスは、たとえば「ａ@ｘ．ｃｏｍ」または「ａ@１．２．３．４」など、ユーザ名とドメイン名またはＩＰアドレス（ＤＮ／ＩＰ）からなる。どちらの例でも、ユーザ名は「ａ」である。

【0022】

ユニフォームリソースロケータ（ＵＲＬ）は、一般に「ｓｅｒｖｉｃｅ－ｎａｍｅ：ＦＱＤＮ／ＩＰ／ｕｒｌ－ｐａｔｈ」という形態のものである。たとえば、「http://www.microsoft.com/windows/help.htm」はＵＲＬである。「ｈｔｔｐ」という部分がサービス名である。「ｗｗｗ．ｍｉｃｒｏｓｏｆｔ．ｃｏｍ」という部分はＦＱＤＮであり、「ｗｉｎｄｏｗｓ／ｈｅｌｐ．ｈｔｍ」はＵＲＬパスである。これはＵＲＬを幾分簡単にしたものであるが、本考察には十分である。

【0023】

次に図１を参照すると、本発明の一態様による特徴抽出およびトレーニングシステム１００の全体的なブロック図が示されている。特徴抽出およびトレーニングシステム１００は、着信メッセージ１１０を処理し、そのメッセージからデータまたは特徴を抽出するものである。そのような特徴は、メッセージおよび／またはその変形の形態において提供される送信元および／または宛先情報の少なくとも一部分から抽出することができる。具体的には、１つまたは複数の着信メッセージ１１０は、メッセージ受信構成要素１２０を介してシステム１００によって受信することができる。メッセージ受信構成要素１２０は、たとえば、着信メッセージ１１０を受信するために、電子メールサーバまたはメッセージサーバ上に位置することができる。一部のメッセージ（たとえば、少なくとも１つ）は既存のフィルタ（たとえば、スパムフィルタ、ジャンクメールフィルタ、ペアレンタル制御フィルタ）に弱く、したがって、ゴミ箱またはジャンクメールフォルダに転送される可能性があるが、送信元および／または宛先データの少なくとも一部分は、機械学習システムと共に、あるいは特徴リストに配置するに際して使用するために、抽出し明瞭にすることができる。

【0024】

メッセージ受信構成要素１２０は、着信メッセージ、またはそのサブセットを特徴抽出構成要素１３０に渡すことができる。特徴抽出構成要素１３０は、フィルタトレーニング、最終的にはスパム検出を容易にするための特徴セットを生成するために、それぞれのメッセージ１１０からデータを抽出することができる。メッセージから抽出されたこのデータまたは特徴は、メッセージ内で見つけられた、および／または埋め込まれた送信元および／または宛先情報に関連する。データまたは特徴の例には、発信元（ｒｅｃｅｉｖｅｄ－ｆｒｏｍ）ＩＰアドレス、返信先（ｒｅｐｌｙ－ｔｏ）電子メールアドレス、ｃｃ：（たとえば、カーボンコピー）電子メールアドレス、各種ＵＲＬ（テキストをベースとするリンク、画像をベースとするリンク、およびテキスト形態のＵＲＬまたはその一部分）、非無料電話番号（たとえば、特にエリアコード）、無料電話番号、ｍａｉｌｔｏ：電子メールアドレスリンク、テキスト形態の電子メールアドレス、ＳＭＴＰ　ＨＥＬＯコマンド内のＦＱＤＮ、ＳＭＴＰ　ＭＡＩＬ　ＦＲＯＭアドレス／リターンパスアドレス、および／または上記のいずれかの少なくとも一部分が含まれる。

【0025】

特徴抽出構成要素１３０は、任意の適切な数のプロセスを実行し、後に機械学習システムで使用するために、メッセージ１１０から様々な特徴セットを抽出することができる。加えて、またはこれに代えて、特徴セットは、他のフィルタトレーニング技法用のリストに配置させるために使用することができる。

【0026】

たとえば、ａ．ｘ．ｃｏｍなどＦＱＤＮは、一般にＩＰアドレスと呼ばれる番号に変換することができる。ＩＰアドレスは、一般に、４つの番号ブロックを備えたドット付きの１０進フォーマットとされる。各ブロックは、点または小数点によって分離され、各番号

ブロックは、０から２５５の範囲に及ぶことができ、番号の各変化は、異なるインターネット名に対応する。たとえば、ａ．ｘ．ｃｏｍは１２３．１２４．１２５．１２６に変換される可能性があり、一方、１２１．１２４．１２５．１２６はｑｒｓｔｕｖ．ｃｏｍを表す可能性がある。番号は、単語ほど容易に認識または記憶することができないため、ＩＰアドレスは通常、それぞれのＦＱＤＮによって参照される。また、ドット付き１０進フォーマットの同じＩＰアドレスを、以下で説明されることになる代替フォーマットで表すこともできる。

【0027】

本発明の一態様によれば、特徴抽出構成要素１３０は、メッセージ１１０内に含まれる発信元ＩＰアドレスに集中して処理することができる。発信元ＩＰアドレスは、少なくとも部分的には、発信元ＩＰ情報に基づくものである。一般に、インターネットを介して送信されたメールは、サーバからサーバに移送され、時に２つのサーバ（たとえば、送信側と受信側）を必要とするだけである。まれではあるが、クライアントがサーバに直接送信することができる。場合によっては、たとえば、ファイアウォールがあるために、メールまたはメッセージが、あるサーバから別のサーバに送信される。具体的には、一部のサーバは、ファイアウォールの内側に位置する可能性があり、したがって、そのファイアウォールの他方の側の指定されたサーバと通信することができるだけである。これは、送信側から受信側に到達するためにメッセージが取るホップの数の増加を引き起こす。そのＩＰアドレスを含む発信元の行により、メッセージがどこから来たか突き止めるため、メッセージのパスを追跡することが容易になる。

【0028】

メッセージ１１０がサーバからサーバへ移動するにつれ、途中の各サーバは、伝送されているサーバの主張しているＦＱＤＮの名前に加えて、メッセージが届けられたＩＰアドレスの識別をメッセージの発信元フィールド（すなわち、「Ｒｅｃｅｉｖｅｄ：」フィールド）の先頭に付加する。このＦＱＤＮは、ＳＭＴＰプロトコルのＨＥＬＯコマンドを介して送信側サーバによって受信側サーバに伝送され、したがって、送信側サーバがそのシステムの外側にある場合には信頼することができない。たとえば、メッセージは、５つのＩＰアドレスとＦＱＤＮが付加された５つの発信元行を有し、したがって、行が付加された逆順（すなわち、最後が先頭）の状態で、メッセージが６つの異なるサーバを通った（すなわち、５回通過した）ことを示すことができる。しかし、各サーバは、より下方の（より早く付加された）行を修正することができる。これは、特にメッセージが複数のサーバ間を移動してきたとき、特に問題をはらむこととなる。各中間サーバは、より早く書き込まれた任意の（より下方の）発信元の行を変えることが可能であるため、スパム送信者は、発信元ＩＰ情報またはスパムメッセージの送信者を偽装するため、メッセージの発信元の行に偽のＩＰアドレスを付加することができる。たとえば、スパムメッセージは、当初、ｔｒｕｓｔｅｄｄｏｍａｉｎ．ｃｏｍ（信頼できるドメイン）から送信されたものであるかのように見え、したがって、メッセージの真の送信源について受信者に偽りを伝える可能性がある。

【0029】

スパムソフトウェアにとって、システムの内側のサーバに送信したシステム外側のＩＰアドレスを容易に識別することは重要である。このＩＰアドレスは、システム内側の受信サーバによって書き込まれるため、正しいＩＰアドレスとして信頼され得る。システム外側の他のＩＰアドレスはすべて、システム外側のサーバによって書き込まれており、したがって、修正されているおそれがあるため、信頼されない可能性がある。受信者システムへのパス内に送信サーバの多数のＩＰアドレスが含まれる可能性があるが、１つを信頼することができるにすぎないため、この信頼できるＩＰアドレスを「送信者の」ＩＰアドレスと呼ぶものとする。

【0030】

スパムフィルタリングソフトウェアがこの送信者のＩＰアドレスを見つけるための１つの方法としては、システムでのメールサーバ構成を知ることが挙げられる。一般に、どの

状況でどの装置がどの他の装置に渡すか知っている場合、送信者のＩＰアドレスを決定することができる。しかし、特に電子メールクライアントにインストールされたスパムフィルタリングソフトウェアにとって、サーバ構成を記述することは都合が悪い場合がある。これに代わる手法には、ＭＸレコードを使用してメッセージの真の送信源を決定するステップが含まれる。ＭＸレコードは、各ドメイン名ごとに、そのドメインについて電子メールの受信者のＦＱＤＮをリストする。そのドメインのＭＸレコード内のエントリに対応するＦＱＤＮに対応するＩＰアドレスが見つかるまで、発信元リストを遡って追跡することができる。この装置が受信した発信元のＩＰアドレスは、送信者のＩＰアドレスである。１．２．３．１０１がｘ．ｃｏｍについて唯一のＭＸレコードであると想像してみる。次いで、１．２．３．１０１から届けられた行を見つけることにより、次の行がｘ．ｃｏｍの着信メールサーバに対応すること、したがって、その行内のＩＰアドレスが、ｘ．ｃｏｍに送信したＩＰアドレスに対応することを知ることができる。

【0031】

下記の表は、メッセージの真の送信源を決定する上記で説明した例示的な解析を示す。

【0032】

【表1】

行	解釈
Received: from a.x.com ([1.2.3.100]) by b.x.com Tue, 22 Apr 2003 13:11:48 -0700	x.comの内部
Received: from mailserver.x.com ([1.2.3.101]) by b.x.com Tue, 22 Apr 2003 12:11:48 -0700	1.2.3.101は、x.comについてのMXレコードであり、それにより、次の行がx.com内部の最初であることがわかる。
Received: from outside.com ([4.5.6.7]) by mailserver.x.com Tue, 22 Apr 2003 11:11:48 -0700	これは、x.comがメッセージを受信した所である。すなわち、これは最後の信頼できる行である。4.5.6.7を送信者のIPアドレスとして使用する。
Received: from trustedsender.com ([8.9.10.11]) by outside.com Tue, 22 Apr 2003 10:11:48 -0700	この行は、4.5.6.7でサーバによって構築された偽物である可能性がある。

【0033】

現在、発信メールサーバをリストするための受け入れられている標準はなく、たとえば、システム内部のＩＰアドレスがシステム外部のＩＰアドレスと異なる場合、あるいは、システムが、ＭＸレコード内にリストされているある装置からＭＸレコード内にリストされている別の装置に間接的に送信する場合、このヒューリスティックは失敗する可能性がある。さらに、ＭＸレコード内のある装置がＭＸレコード内の別の装置に送信した場合に発生する可能性があるように、上記のように見つけられた送信者のＩＰがそのシステムの内部にあると判明した特別な場合、プロセスは上記のように継続される。さらに、ある種のＩＰアドレスを内部として検出することができる（というのは、内部ＩＰアドレスのためだけに使用されている形態である、１０．ｘ．ｙ．ｚまたは１７２．１６．ｙ．ｚから１７２．３１．ｙ．ｚまたは１９２．１６８．０．ｚから１９２．１６８．２５５．ｚの形態のものであるからである）。すなわち、システム内部のどのアドレスも信頼することができる。最後に、発信元の行が「Ｒｅｃｅｉｖｅｄ　ｆｒｏｍ　ａ．ｘ．ｃｏｍ［１．２．３．１００］」という形態のものであり、ａ．ｘ．ｃｏｍのＩＰアドレスルックアップが１．２．３．１００を生じる、あるいは、１．２．３．１００の逆ＩＰアドレスルックアップがａ．ｘ．ｃｏｍを生じる場合、また、ｘ．ｃｏｍがそのシステムである場合には、次の行もまた信頼することができる。

【0034】

これらの観察結果を使用して、送信者のＩＰアドレスを見つけられることが多い。例示的な疑似コードは次の通りである。

【0035】

【表2】

```
bool fFoundHostInMX;
if (external IP address of MX records matches internal IP
address of MX records)
{
```

【0036】

【表3】

```
    fFoundHostInMX = FALSE; # it's worth looking for
} else {
    fFoundHostInMX = TRUE; # it's not worth looking for,
pretend we already found it
}

for each received from line of the form Received from a.b.c
[i.j.k.l] {
    if i.j.k.l in MX records of receiver domain
    {
        fFoundHostInMX = TRUE;
        continue;
    }
    if not fFoundHostInMX
    {
        # Has not yet gone through an MX record, must be
internal
        continue;
    }
    if i.j.k.l is of form
        10.x.y.z or
        172.16.y.z to 172.31.y.z or
        192.168.0.z to 192.168.255.z
    {
        # Must be internal
        continue;
    }
    if DNS lookup of a.b.c yields i.j.k.l and b.c is
receiver domain
    {
        # Must be internal
        continue;
```

【0037】

【表4】

```
    }
    Output sender's alleged FQDN a.b.c and sender's actual
IP address i.j.k.k
}
If we reach here, then Error: unable to identify sender's
alleged FQDN and sender's actual IP address
```

【0038】
他の送信元および宛先機能の場合と同様に、送信者のＩＰアドレスを用いて多くの処理を行うことができる。第１に、ブラックリストと呼ばれることもある一様に悪質な送信者のリストに追加することができる。ブラックリストは、信頼できないメッセージをフィルタし、遮断し、または、それらをさらに調査することができる適切なフォルダまたは位置に向けて送り直すために後に使用することができる。

【0039】
また、他の型のリストを生成し、クライアントベースのアーキテクチャとサーバベースのアーキテクチャのどちらでもフィルタとして実行することができる。クライアントアーキテクチャでは、ユーザは（たとえば、メーリングリスト、個人など）誰からのメールを受信するかをクライアント電子メールソフトウェアに通知することができる。信頼される電子メールアドレスに対応するレコードのリストを、ユーザが手動で、または自動的に生成することができる。したがって、電子メールアドレス「ｂ＠ｚｙｘ．ｃｏｍ」を有する送信者がユーザに電子メールメッセージを送信すると考える。送信者の電子メールアドレスｂ＠ｚｙｘ．ｃｏｍは、ユーザ名「ｂ」とＦＱＤＮ／ＩＰ「ｚｙｘ．ｃｏｍ」を含む。クライアントが送信者（ｂ＠ｚｙｘ．ｃｏｍ）から着信メッセージ１１０を受信すると、そのユーザの電子メールアドレスについて信頼される送信者リストを探索し、「ｂ＠ｚｙｘ．ｃｏｍ」が有効かつ信頼されるアドレスであることをユーザが示しているかどうか判定することができる。サーバアーキテクチャの場合、リストは、サーバ上に直接配置することができる。したがって、メッセージがメッセージサーバに到着したとき、それぞれの特徴（たとえば、送信者のＩＰアドレス、ＭＡＩＬ　ＦＲＯＭまたはＨＥＬＯフィールド内のドメイン名、ならびに他の送信元および／または宛先情報）をメッセージサーバ上に配置されるリストと比較することができる。有効な送信者からのものと決定されたメッセージは、クライアントベースの送達プロトコルまたはサーバベースの送達プロトコルに従って、意図された受信者に送達することができる。しかし、疑わしいまたは悪質な特徴のリスト内の送信元または宛先の特徴を含むと決定されたメッセージは、廃棄するため、または他の方法で特別に処理するため、スパムまたはジャンクメールフォルダに移動することができる。

【0040】
信頼される送信元の特徴または悪質な送信元の特徴のリストに配置することに代わる方法として、送信者の送信元の特徴（たとえば、ＩＰアドレス、主張されている発信元アドレス）を１つまたは複数の特徴として抽出し、フィルタ構築および／またはトレーニングのために機械学習技法と共に後に使用することができる。

【0041】
ＩＰアドレスは、メッセージヘッダの任意の一部内の電子メールアドレス（たとえば、送信者のアドレスまたは返信先アドレス内のＦＱＤＮ上のＩＰルックアップ）から、または、メッセージの本文内に埋め込まれたＵＲＬリンクのドメイン名部分のＩＰアドレスルックアップから、またはＩＰアドレスがＵＲＬのＦＱＤＮ／ＩＰ部分として見られる場合には直接そこから導出することができる。さらに、後に説明するように、ＩＰアドレスはいくつかの属性を有し、その各々を、機械学習システムの特徴として、またはユーザによ

ってポピュレートされるリストの要素として使用することができる。したがって、第２の手法では、特徴抽出構成要素１３０は、ＩＰアドレスの多数の下位区分を利用し、追加の特徴を生成することができる。

【0042】

上記の特徴の任意の組合せは、各着信メッセージ１１０から抽出することができる。典型的にはメッセージすべてを使用することができるが、メッセージは、ランダムに、自動的に、および／または手動で選択し、特徴抽出に加わることができる。抽出された特徴セットは、機械学習システム、あるいはスパムフィルタなどフィルタ１５０を構築および／またはトレーニングする任意の他のシステムなど、フィルタトレーニング構成要素１４０に後に加えられる。

【0043】

図２を参照すると、本発明の一態様による着信メッセージ２１０の１つまたは複数の特徴を明瞭化（ｄｅｏｂｆｕｓｃａｔｅ）または正規化することを容易にする特徴抽出システム２００が示されている。最終的には、少なくとも部分的には正規化された特徴の１つまたは複数に基づいて、フィルタを構築することができる。システム２００は、たとえば、図のように直接、またはメッセージ受信側（図１）によって間接的に着信メッセージ２１０を受信する特徴抽出構成要素２２０を含む。特徴抽出のために選択された、または特徴抽出に加わる着信メッセージは、ユーザ嗜好に従ってシステム２００の対象とすることができる。またはこれに代えて、実質的にすべての着信メッセージを特徴抽出のために使用可能にし、また特徴抽出に加えることができる。

【0044】

特徴抽出は、メッセージ２１０からの送信元および／または宛先情報に関連付けられた１つまたは複数の特徴２３０（特徴$_1$ ２３２、特徴$_2$ ２３４、特徴$_M$ ２３６とも呼ばれ、ただし、Ｍは１以上の整数である）を引き出すことが必要である。送信元情報は、メッセージの送信者を示す要素、ならびにサーバドメイン名に関連し、およびメッセージが来た所を指定する識別情報に関連する可能性がある。宛先情報は、受信者がそのメッセージに対する応答を誰に、またはどこに送信することができるかを示すメッセージの要素に関連する可能性がある。送信元および宛先情報は、メッセージ受信者に見える、または見えない状態で（たとえば、テキストとして、または画像内に埋め込まれて）、メッセージのヘッダ内に、ならびにメッセージの本文内に見いだすことができる。

【0045】

スパム送信者は、従来のスパムフィルタによって検出されるのを回避するために正体を偽装し、および／または不明瞭化しようとする傾向がよくあるので、システム２００は、１つまたは複数の抽出された特徴２３０、またはその少なくとも一部分を明瞭化するようにすることを促進する特徴正規化構成要素２４０を備える。特徴正規化構成要素２４０は、抽出された特徴２３０（たとえば、ＦＱＤＮであるが、これは、ブロックおよびＭＸレコードのディレクトリを調べ、および／またはその現行フォーマットに従って変換される）を解析し、次いでそれらと、たとえば既存のスパム送信者リスト、非スパム送信者リスト、および／またはペアレンタル制御リストのデータベースとを比較することなどにより、抽出された特徴２３０を処理および／または分析することができる。図４において上記で説明したいくつかの場合には、抽出された特徴がＵＲＬであるときなど、接頭語および／または接尾語を除去して特徴を正規化すること、およびそのＵＲＬがスパム送信者のウェブサイトを指しているか、あるいは本物の送信源を指しているかを識別することを容易にすることもできる。

【0046】

特徴が正規化された後に、機械学習システムなどトレーニングシステム２６０がそれらの少なくともサブセット２５０を使用し、フィルタ２７０を構築および／または更新することができる。フィルタは、たとえばスパムフィルタおよび／またはジャンクメールフィルタとして使用するためにトレーニングすることができる。さらに、非スパム源（たとえば、送信者の発信元電子メールアドレス、送信者のＩＰアドレス、埋込み電話番号、およ

び／またはＵＲＬ）および／または非スパム送信者を示すものなどの肯定的な特徴によって、ならびにスパム送信者を識別したり、スパム送信者に関連付けられたりするものなどの否定的な特徴によってフィルタを構築および／またはトレーニングすることができる。

【0047】

またはこれに代えて、あるいはこれに加えて、特徴セットは、新しいスパム特徴リスト２８０に配置するために、または既存のスパム特徴リスト２８０に追加するために使用することができる。また、問題のないアドレスのリスト、悪質なアドレスのリスト、問題のないＵＲＬのリスト、悪質なＵＲＬのリスト、問題のない電話番号のリスト、悪質な電話番号のリストなど、特定の抽出された特徴に対応するように他のリストを生成することができる。問題のない特徴リストは、非スパム送信者、履歴からみて本物の送信者、および／または非スパム送信者である尤度がより高い送信者（例えば、スパム源でない公算が９０％以下）を識別することができる。逆に、悪質な特徴リストは、スパム送信者、潜在的なスパム送信者、および／またはスパムである尤度が比較的高い送信者（たとえば、９０％以下のスパム源）に対応付けすることができる。

【0048】

次に図３～６は、本発明のいくつかの態様に従ってスパム検出および防止を容易にするため、ＩＰアドレス、ＦＱＤＮ、電子メールアドレス、およびＵＲＬからそれぞれ導出または抽出することができる例示的な特徴を示す。

【0049】

図３は、本発明の態様によるＩＰアドレス３００の例示的な内容を示す。ＩＰアドレス３００は３２ビット長であり、ドット付き１０進フォーマット（たとえば、それぞれ３桁までの４ブロックであり、各ブロックはピリオドによって分離され、３桁の各ブロックは、０と２５５の間で可分の任意の数である）で表されるとすると、ブロック（たとえば、ネットブロック）内に割り振られる。ブロックは、クラスＡ、クラスＢ、クラスＣなど、クラスに割り当てられる。各ブロックは、ＩＰアドレスの複数のセットを含み、ブロック当たりのＩＰアドレス数は、クラスによって変わる。すなわち、クラス（すなわち、Ａ、Ｂ、またはＣ）に応じて、ブロック当たり、より多くの、またはより少ないアドレスが割り当てられる可能性がある。ブロックサイズは、通常、２のべき乗であり、同じブロック内のＩＰアドレスのセットは、最初のｋ個の２進数字を共有し、最後の３２－ｋ（３２引くｋ）個の２進数字が異なることになる。したがって、各ブロックは、その共有されている最初のｋ個のビットに従って識別することができる（ブロックＩＤ３０２）。特定のＩＰアドレス３００に関連付けられたブロックＩＤ３０２を決定するために、ユーザは、ａｒｉｎ．ｎｅｔなどブロックのディレクトリを調べることができる。さらに、ブロックＩＤ３０２は、特徴として抽出および使用することができる。

【0050】

しかし、場合によっては、ブロック内のＩＰアドレスのグループが分割されて処分され、また任意の回数だけ使いまわされる可能性があるため、ａｒｉｎ．ｎｅｔを参照してもブロックＩＤ３０２を容易に決定することができない。そのような場合には、ユーザまたは抽出システムは、各々のＩＰアドレスについて１回または複数回、ブロックＩＤ３０２を推測することができる。たとえば、ユーザは、機械学習システムによって後に使用するための別の特徴として、および／または特徴リスト（たとえば、問題のない特徴リスト、スパム特徴リストなど）上の要素として、少なくとも最初の１ビット３０４、少なくとも最初の２ビット３０６、少なくとも最初の３ビット３０８、少なくとも最初のＭビット３１０（すなわち、Ｍは１以上の整数）、および／または少なくとも最初の３１ビット３１２まで抽出することができる。

【0051】

実際には、たとえば、ＩＰアドレスの最初の１ビットを特徴として抽出および使用し、ＩＰアドレスがスパム送信者を指しているか、あるいは非スパム送信者を指しているか判定することができる。他のメッセージから抽出された他のＩＰアドレスからの最初の１ビットを比較し、少なくとも１つのブロックＩＤを決定することを容易にすることができる

のである。次いで、少なくとも１つのブロックを識別することは、そのメッセージがスパム送信者からのものかどうか見分けるのに役立つ可能性がある。さらに、最初のＭビットを共有するＩＰアドレス群をそれらの他の抽出された特徴について比較し、ＩＰアドレスが本物の送信者からのものかどうか、および／またはメッセージがそれぞれスパムかどうか突き止めることができる。

【0052】

また、ＩＰアドレスを階層（３１４）によって構成することができる。すなわち、１組のより高次のビットを特定の国に割り振ることができる。その国では、ＩＳＰ（インターネットサービスプロバイダ）を一定のサブセットに割り振り、次いで、そのＩＳＰでは、特定の会社を一定のサブセットに割り振ることができる。したがって、同じＩＰアドレスでも様々なレベルをもつことに意義がある。たとえば、ＩＰアドレスが韓国に割り振られたブロックから来たという事実は、そのＩＰアドレスがスパム送信者に関連付けられているかどうか判定する上で有用である可能性がある。そのＩＰアドレスが、スパム送信者に対して厳しいポリシーを有するＩＳＰに割り振られたブロックの一部である場合、これもまた、そのＩＰアドレスがスパム送信者に割り当てられていないと決定する上で有用である可能性がある。したがって、ＩＰアドレスの最初の１～３１ビットの各々を、ＩＰアドレス群の少なくともサブセットの階層構成３１４と組み合わせて使用することにより、ユーザは、ＩＰアドレスが割り振られた方法を実際に知ることなしに（たとえば、ブロックＩＤを知ることなしに）、様々なレベルで自動的に情報を学習することができる。

【0053】

上記で説明した特徴に加えて、特徴の稀少性３１６（たとえば、特徴の発生があまり普通でない）は、たとえば、適切な計算を実行すること、および／または着信メッセージをサンプリングする際に特徴が現れる周期またはカウントを比較する統計データを使用することによって決定することができる。実際には、めったにないＩＰアドレス３００が、電子メールを送達するために使用されるダイヤルアップ回線である可能性があるが、これはスパム送信者によってしばしば使用される戦法である。スパム送信者は、その識別および／または位置を頻繁に修正する傾向がある。したがって、その特徴は普通である、またはめったにないという事実は有用な情報となる。したがって、特徴の稀少性３１６は、機械学習システムの特徴として、および／または少なくとも１つのリスト（たとえば、稀少特徴リスト）の一部として使用することができる。

【0054】

図４は、たとえばｂ．ｘ．ｃｏｍなどのＦＱＤＮ４００の例示的な特徴の内容を示す。ＦＱＤＮ４００は、たとえば、ＨＥＬＯフィールドから抽出することができ（たとえば、送信者であると主張するＦＱＤＮ）、一般に、ホスト名４０２およびドメイン名４０４を含む。ホスト名４０２は、その例によれば「ｂ」である特定のコンピュータを指す。ドメイン名４０４は、インターネット上の少なくとも１つの装置または装置のグループの名前を指す。本例では、「ｘ．ｃｏｍ」は、ドメイン名４０４を表す。ＦＱＤＮ４００の階層の内容は、内容４０６によって表される。具体的には、Ｂ．Ｘ．ＣＯＭ４０８（完全なＦＱＤＮ４００）は、部分的にＸ．ＣＯＭ４１０（部分的なＦＱＤＮ）に分解することができ、次いで、これをＣＯＭ４１２（部分的なＦＱＤＮ）に分解することができ、それにより、各部分的ＦＱＤＮを特徴として使用することができる。

【0055】

発信元情報など、いくつかの特徴は、主にＩＰアドレスとして存在する。したがって、ＦＱＤＮ４００を、（図３に示すように）ＩＰアドレス３００に変換し、追加の特徴として分析することは有用である可能性がある。というのは、新しいホスト名やドメイン名を作成するのは比較的容易であるが、新しいＩＰアドレスを得るのは比較的困難だからである。

【0056】

残念ながら、ドメインの所有者は、明らかに異なる装置すべてを同じ場所にマッピングさせることができる。たとえば、「ａ．ｘ．ｃｏｍ」という名前の装置の所有者は、「ｘ

．ｃｏｍ」の同じ所有者である「ｂ．ｘ．ｃｏｍ」の所有者と同じである可能性がある。したがって、スパム送信者は、そのメッセージが、ドメイン４０４「ｘ．ｃｏｍ」からのものではなく、ＦＱＤＮ４００「ｂ．ｘ．ｃｏｍ」からのものであると信じるように従来型フィルタを容易に欺き、それにより、実際には、そのメッセージがスパムであった、またはスパムである可能性がより高いことをドメイン４０４「ｘ．ｃｏｍ」が示していたかもしれない場合でも、メッセージがスパムフィルタを通過させてしまう可能性がある。したがって、メッセージの送信元および／または宛先情報を抽出するとき、アドレスを単にドメイン名４０４に分解することは有用である。またはこれに代えて、またはこれに加えて、完全ＦＱＤＮ４００を特徴として抽出することができる。

【0057】

システムによっては、ペアレンタル制御システムなど追加の資源が使用可能な場合がある。これらの資源は、しばしば、ポルノや暴力など「型」または質的評価をホスト名、および／またはＵＲＬに割り当てることができる。抽出された特徴は、そのような資源を使用して、型によってさらに分類することができる。次いで、改良されたスパム関連フィルタを構築すること、および／またはトレーニングすることと併せて、特徴の型４１４を追加の特徴として使用することができる。またはこれに代えて、先に識別されている様々な特徴の型に対応して、リストを生成することができる。特徴の型４１４には、それだけには限らないが、性またはポルノ関連の特徴、人種および／または憎悪発言関連の特徴、肉体強化の特徴、収入または金銭的解決策の特徴、住宅購入の特徴などが含まれ、これらは、メッセージの一般的な目的を識別する。

【0058】

最後に、特徴の稀少性３１６、または特徴の型の希少性（上記図３参照）を、図３において上記で説明した別の特徴とすることができる。たとえば、ＦＱＤＮ４００「ｂ．ｘ．ｃｏｍ」のホスト名「Ｂ」４０２など、メッセージから抽出された特徴を特徴の型に共通する事例、例えばポルノ題材とすることができる。したがって、この特徴がメッセージから抽出され、次いでポルノ題材特徴リスト上で見つかったときには、そのメッセージは、スパムである可能性がより高い、またはあらゆる年齢にとってふさわしくない／不適切である、若しくは成人向け内容（たとえば、成人レーティング）を構成する、などと結論を下すことができる。したがって、各リストは、その特定の型の、より共通（ｃｏｍｍｏｎ）の特徴を含むことができる。またはこれに代えて、一般に、対応するＩＰアドレスがスパムメッセージ内で共通して見出され、したがってスパムの共通の特徴として指定することができる。さらに、特徴の共通性（ｃｏｍｍｏｎａｌｉｔｙ）および／または希少性を、装置学習または他の規則をベースとするシステム用の別の特徴として使用することができる。

【0059】

図５は、ＦＱＤＮ４００ならびにユーザ名５０２などいくつかの追加の特徴を含む、ａ＠ｂ．ｘ．ｃｏｍという電子メールアドレス５００の例示的な特徴の内容を示す。電子メールアドレス５００は、メッセージのＦｒｏｍフィールド、ｃｃ（カーボンコピー）フィールド、ｒｅｐｌｙ－ｔｏフィールドから、ならびに、メッセージ本文内のｍａｉｌｔｏ：リンクのいずれか（たとえば、ｍａｉｌｔｏ：リンクは、クリックされたとき特定のアドレスへのメールを生成する特別な種類のリンクである）から、さらに使用可能な場合、ＳＭＴＰプロトコル内で使用されるＭＡＩＬ　ＦＲＯＭコマンドから抽出することができる。また、電子メールアドレス５００は、メッセージの本文内にテキストとして埋め込むことができる。場合によっては、メッセージ内容は、そのメッセージに応答するとき「ｒｅｐｌｙ　ａｌｌ（全員に返信）」機能を使用するように受信者を導くものである可能性がある。そのような場合には、ｃｃフィールド内のアドレス、および／または「ｔｏ」フィールド内に含まれるアドレスの少なくとも一部分（複数の受信者がリストされている場合）もまた、返信先であろう。したがって、これらのアドレスのそれぞれを１つまたは複数の特徴として抽出し、スパム送信者識別／防止を容易にすることができるであろう。

【0060】

「ａ@ｂ．ｘ．ｃｏｍ」という電子メールアドレス５００は、様々な要素または下位区分に分析することができ、これらの要素もまた、特徴として抽出および使用することができる。具体的には、電子メールアドレスは、ユーザ名５０２と、さらに追加の特徴に細分化することができるＦＱＤＮ５０４（たとえば、図４のＦＱＤＮ４００参照）とを含む。使用、認識、および想起の容易さなど、いくつかの実際的な理由により、電子メールアドレスは、通常、ＩＰアドレスではなくＦＱＤＮを使用して表記される。

【0061】

この例では、「ａ@ｂ．ｘ．ｃｏｍ」は、ユーザ名５０２「ａ」を含む。したがって、「ａ」を１つの特徴として抽出することができる。同様に、ＦＱＤＮ５０４「ｂ．ｘ．ｃｏｍ」を、少なくとも１つの他の特徴として電子メールアドレスから抽出することができる。電子メールアドレス５００のＦＱＤＮ５０４部分は、図４において上記により詳しく説明した特徴の型４１４を決定することを容易にするため、ペアレンタル制御フィルタに通すことができる。したがって、電子メールアドレス５００のＦＱＤＮ部分に関係する特徴の型を追加の特徴として使用することができる。

【0062】

電子メールアドレスに加えて、スパム送信者は、ＵＲＬを介してアクセスを受けることがよくある。図６は、本発明の態様による例示的なＵＲＬ６００（たとえば、ｘ．ｙ．ｃｏｍ／ａ／ｂ／ｃ）と、そこから抽出された複数の特徴とを示す。ＵＲＬ６００は、メッセージ本文内のテキストとして、および／またはメッセージ本文内の画像として埋め込まれる可能性がある。たとえば、スパムメッセージは、ウェブサイトに対するポインタを含み、それにより、受信者をスパム送信者のウェブサイトまたは関連サイトに導く可能性がある。

【0063】

ＵＲＬは、ＩＰアドレスについてしたと同様な方法で明瞭化することができる。ＵＲＬ６００を明瞭化する前に、まず、たとえば、ｈｔｔｐ：／／、ｈｔｔｐｓ：／／、ｆｔｐ：／／、ｔｅｌｎｅｔ：／／など任意の接頭語（たとえば、サービス名）を除去することができる。さらに、「@」記号（たとえば、１６進表記で％４０）がＵＲＬの中間に現れた場合、接頭語（たとえば、ｈｔｔｐ：／／）と「@」記号の間にどんなものがあっても除去することができ、それからＵＲＬ６００を正規化することができる。スパム送信者による別の戦法または策略の形態として、接頭語と「@」記号の間にテキストを組み込むことにより、受信者が誘導されている真のページ位置に関してメッセージ受信者を混乱させることも考えられる。

【0064】

たとえば、http://www.amazon.com@121.122.123.124/info.htmは、メッセージ受信者にとって、このページがｗｗｗ．ａｍａｚｏｎ．ｃｏｍであるかのように見える。したがって、受信者がそのリンクを、およびより重要なことには、そのメッセージ送信者を信頼する傾向がより強くなる可能性がある。それに反して、真のページ位置は、実際にスパム関連ウェブページに対応する「１２１．１２２．１２３．１２４」にある。しかし、場合によっては、自動ログインを容易にするために、本物の送信者が、ＵＲＬ６００のこの部分にログイン名およびパスワードなど認証情報を組み込む可能性がある。

【0065】

正規化および明瞭化した後に、ＵＲＬ６００は、本質的にｘ．ｙ．ｃｏｍ／ａ／ｂ／ｃとして表すことができ、ただし、ｘ．ｙ．ｃｏｍ６３０は装置の名前（ＦＱＤＮ）であり、ａ／ｂ／ｃ（たとえば、接尾語）はその装置上のファイルの位置である。ｘ．ｙ．ｃｏｍ／ａ／ｂ／ｃ６００がスパム送信者を識別する場合には、ｘ．ｙ．ｃｏｍ／ａ／ｂ６１０およびｘ．ｙ．ｃｏｍ／ａ６２０もまた、同じまたは関連するスパム送信者を識別する可能性が非常に高い。したがって、ＵＲＬ６００の末端部分またはパスウェイは１度に１つの部分が分解され、たとえば、機械学習システムまたはリスト用の追加の特徴を得ることができる。これにより、すべてが実際にパターンに気付かれないような方法でスパム送信者につながる多数の様々な配置をスパム送信者が作成するのはより困難になる。

【0066】
接尾語が分解されたとき、図４において上記で説明したように、ＦＱＤＮ６３０をさらに構文解析し、追加の特徴を得ることができる。さらに、上記で図３において示されているように、ＦＱＤＮ６３０もまた、ＩＰアドレスに変換することができる。したがって、ＩＰアドレスに関連する様々な特徴もまた、特徴として使用することができる。
【0067】
いくつかのＵＲＬは、ｎｎｎ．ｎｎｎ．ｎｎｎ．ｎｎｎ／ａ／ｂ／ｃなど、ＦＱＤＮでなくＩＰアドレス（たとえば、ドット付き１０進フォーマット）で記述される。接尾語は「ｃ」で始まる連続する順番で除去することができ、各段階で、得られる（部分的な）ＵＲＬを特徴として使用することができる（たとえば、ｎｎｎ．ｎｎｎ．ｎｎｎ．ｎｎｎ／ａ／ｂ、ｎｎｎ．ｎｎｎ．ｎｎｎ．ｎｎｎ／ａ、ｎｎｎ．ｎｎｎ．ｎｎｎ．ｎｎｎは、すべてドット付き１０進フォーマットのＵＲＬから抽出することが可能な特徴である）。引き続いて、（たとえば、接頭語と接尾語がない）ＩＰアドレスを特徴として使用することができる。次いで、そのＩＰアドレスをそのネットブロックにマップすることができる。ネットブロックが確認できない場合には、ＩＰアドレスの最初の１、２～最初の３１ビットまでの各々を別の特徴として使用し、複数回推測することができる（図３参照）。
【0068】
ドット付き１０進フォーマットに加えて、ＩＰアドレスは、ｄｗｏｒｄ（ダブルワード）フォーマット（たとえば、それぞれ１０進法で２つの１６ビットからなる２進ワード）、８進フォーマット（たとえば、８進法）、１６進フォーマット（たとえば、１６進法）で表すことができる。実際には、スパム送信者は、たとえば、ドメイン名部分を％ｎｎ表記（ｎｎは１６進数字の対）を使用して符号化することにより、ＩＰアドレス、ＵＲＬ、ｍａｉｌｔｏ：リンク、および／またはＦＱＤＮを不明瞭化する可能性がある。
【0069】
いくつかのＵＲＬは、ユーザを混乱させ、または騙すために使用することができるリダイレクタを含む可能性がある。リダイレクタは、ＵＲＬのＩＰアドレス内で「？」に続くパラメータまたはパラメータのセットであり、別のウェブページに向き直すようにブラウザに指令する。たとえば、ＵＲＬは「www.intendedpage.com?www.actualpage.com.」として現れる可能性があり、ブラウザは実際に「ｗｗｗ．ａｃｔｕａｌｐａｇｅ．ｃｏｍ」を指しており、予想されている「ｗｗｗ．ｉｎｔｅｎｄｅｄｐａｇｅ．ｃｏｍ」ではなくそのページをロードする。したがって、ＵＲＬ内に含まれるパラメータもまた、特徴として抽出するために考慮することができる。
【0070】
次に、本発明による様々な方法について、一連の動作を介して説明する。いくつかの動作は、本発明に従って様々な順序で、および／または、本明細書に示され説明されるものからの他の動作と同時に行うことができるため、本発明は、動作の順序によって制限されないことを理解されたい。たとえば、方法は、これに代えて、状態図など一連の相互に関係のある状態または事象として表すことができることを、当業者なら理解できるであろう。さらに、本発明による方法を実施するのに図に示される動作のすべてが必要とされるわけではない。
【0071】
図７は、本発明の態様によるフィルタをトレーニングすることを容易にする例示的なプロセス７００のフローチャートを示す。プロセス７００は、プロセス７１０でメッセージ（たとえば、少なくとも１つのメッセージ）を受信することで開始することができる。メッセージは、たとえば、サーバによって受信されると、サーバ部の既存のフィルタ（たとえば、スパムフィルタ）は、少なくとも部分的にはフィルタによって既に学習された１組の基準に基づいて、そのメッセージをスパムの可能性があるものと、スパムの可能性がないものとともに分類することができる。プロセス７２０において、メッセージを構文解析し、そこから１つまたは複数の特徴を抽出する。特徴の抽出については、（図１１において下記の）プロセス７２５でさらに詳しく説明する。特徴の例には、ｒｅｃｅｉｖｅｄ　ｆ

ｒｏｍフィールド、ｒｅｐｌｙ－ｔｏフィールド、ｃｃフィールド、ｍａｉｌｔｏ：フィールド、ＭＡＩＬ　ＦＲＯＭ　ＳＭＴＰコマンド、ＨＥＬＯフィールド、テキスト内に、または画像として埋め込まれたＵＲＬアドレス、および／または非無料電話番号（たとえば、地理的領域をマップするためのエリアコード）、ならびにメッセージ本文内のテキスト内に位置する情報（たとえば、送信者のＩＰアドレス）が含まれる。

【0072】

抽出された（および／または正規化された）特徴ならびにメッセージの分類（たとえば、スパムまたは非スパム）を、プロセス７３０でトレーニング用データセットに追加することができる。プロセス７４０で、上記（たとえば、プロセス７１０、プロセス７２０、プロセス７３０）は、実質的にすべての他の着信メッセージについて、それらがそれに応じて処理されるまで繰り返すことができる。プロセス７５０で、有用であると思われる特徴、または最も有用な特徴をトレーニング用セットから選択することができる。そのような選択された特徴を使用し、プロセス７６０で、たとえば機械学習アルゴリズムにより、機械学習フィルタなどフィルタをトレーニングすることができる。

【0073】

機械学習フィルタは、トレーニングされた後に、図８の例示的な方法８００によって説明されるように、スパム検出を容易にするために使用することができる。方法８００は、８１０でメッセージを受信することで開始される。プロセス８２０で、図１１に関連して以下で説明するように、１つまたは複数の特徴がメッセージから抽出される。プロセス８３０で、抽出された特徴が、たとえば、機械学習システムによってトレーニングされたフィルタを通過する。引き続いて、「スパム」「非スパム」などの判定、またはメッセージがスパムである確率が、機械学習システムから得られる。メッセージの内容に関して判定が得られた後に、適切な措置をとることができる。措置の型には、それだけには限らないが、メッセージを削除すること、メッセージを特別なフォルダに移動すること、メッセージを隔離すること、受信者がメッセージにアクセスできるようにすることが含まれる。

【0074】

またはこれに代えて、メッセージから抽出された特徴と共に、リストをベースとする活動を実施することができる。図９は、少なくとも部分的には、抽出された特徴と、スパムまたは非スパム（あるいは、スパムである可能性が高い、または可能性が低い）と分類された受信メッセージ内でのその発生に基づいてリストを構築し、かつリストに配置するための例示的なプロセス９００のフローチャートを示す。プロセス９００は、プロセス９１０で、メッセージを受信することによって開始される。引き続いて、プロセス９２０で、たとえばメッセージ送信者のＩＰアドレスなど、いくつかの注目される特徴が抽出される。メッセージが受信された後のある時点で、メッセージを、たとえば、既存のフィルタによってスパムまたは非スパムと分類することができる。プロセス９３０で、メッセージの分類（たとえば、スパムまたは非スパム）に従って、特徴を増分によって計数することができる。プロセス９４０で、実質的にすべてのメッセージが（たとえば、プロセス９１０、プロセス９２０、プロセス９３０で）処理されるまで繰り返すことができる。その後、プロセス９５０で、特徴のリストを作成することができる。たとえば、９０％問題なし（たとえば、その時点の９０％非スパム、または着信メッセージの９０％で非スパム）である送信者ＩＰアドレスについて、あるリストを作成することができる。同様に、９０％悪質（スパム）である送信者ＩＰアドレスについて別のリストを作成することができる。同様の方法で、他の特徴について他のリストを作成することができる。

【0075】

これらのリストは動的に更新することができることを理解されたい。すなわち、これらのリストは、新しいメッセージの追加のグループが処理されたとき更新することができる。したがって、ある送信者のＩＰアドレスが、当初、問題のないリスト上で見つかり、次いで、その後ある時点で、悪質リスト上で見つかる可能性がある。というのは、一部のスパム送信者が（たとえば、フィルタならびに受信者の「信頼」を得るために）最初に問題のないメールを送信し、次いで、実質的にスパムだけを送信し始めることが普通だからで

ある。

【0076】

これらのリストは、様々な方法で使用することができる。たとえば、フィルタをトレーニングするために機械学習システムによって使用されるトレーニング用セットを生成するために使用することができる。そのようなものは、図１０で次に説明される例示的なプロセス１０００によって示されている。図１０によれば、プロセス１０００は、プロセス１０１０でメッセージを受信することによって始めることができる。メッセージを、たとえば、スパムまたは非スパムと分類することができる。プロセス１０２０で、それだけには限らないが、送信者のＩＰアドレスを含む特徴をメッセージから抽出することができる。プロセス１０３０で、抽出された特徴およびメッセージの分類がトレーニング用セットに追加され、このトレーニング用セットは、後に機械学習システムをトレーニングするために使用される。

【0077】

引き続いてプロセス１０４０で、その送信者ＩＰアドレスが存在する特定のリストに対応する特別な特徴がトレーニング用セット内に含まれる。たとえば、その送信者ＩＰアドレスが「９０％問題なし」リスト上にあるならば、トレーニング用セットに追加された特徴は「９０％問題なしリスト」となるはずである。プロセス１０５０で、先行するステップ（たとえば、プロセス１０１０、プロセス１０２０、プロセス１０３０、プロセス１０４０）は、実質的にすべての着信メッセージが処理されるまで繰り返すことができる。いくつかの特徴は、フィルタをトレーニングする目的にとって他の特徴より有用である可能性があるため、プロセス１０６０で、最も有用な特徴または特徴群が、部分的にはユーザ嗜好に基づいて選択され、機械学習アルゴリズムを使用して、スパムフィルタなどフィルタをトレーニングするために使用される。

【0078】

さらに、たとえば、テストメッセージ、新しいメッセージ、および／または疑わしいメッセージと比較するために、ＩＰアドレスの動的リストを構築することができる。しかし、この例では、ＩＰアドレス自体は特徴ではない。むしろＩＰアドレスの品質が特徴である。またはこれに代えて、またはこれに加えて、リストを他の方法で使用することができる。実際には、たとえば、疑わしいＩＰアドレスのリストは、悪質であるとして送信者にフラグし、それに応じて、それらのメッセージを疑いながら処理するために使用することができる。

【0079】

次に、図１１に参照すると、図7～１０においてそれぞれ上記で説明されるプロセス７００、プロセス８００、プロセス９００、プロセス１０００に関連して、メッセージから特徴を抽出する例示的な方法１１００のフローチャートが示されている。方法１１００は、発信元ＩＰアドレスまたはその一部分を抽出し、プロセス１１１０で正規化して開始することができる。また、プロセス１１１０で、その発信元ＩＰアドレスから追加の特徴を抽出するため、そのＩＰアドレスをビット毎（たとえば、図3で説明しているように、最初の１ビット、最初の２ビット～最初の３１ビットまで）の処理にかけることができる。さらに、主張されている送信者のホスト名もまた、プロセス１１１０で抽出することができる。次に、正規化された発信元ＩＰアドレスおよび送信者ホスト名の特徴群を、機械学習システムまたは関連トレーニング用システムの特徴群として使用することができる。

【0080】

オプションであるが、プロセス１１２０にて、「Ｆｒｏｍ：」行の内容を抽出および／または正規化し、後に特徴として使用することができる。プロセス１１３０で、同様に「ＭＡＩＬ　ＦＲＯＭ　ＳＭＴＰ」コマンドの内容を、特徴として使用するために抽出および／または正規化することができる。

【0081】

次いで、方法１１００の処理を行い、メッセージ内に含まれる可能性のある他の可能な特徴を捜すことができる。たとえば、オプションで、（必要な場合）プロセス１１４０に

てｒｅｐｌｙ－ｔｏフィールド内の内容を抽出および正規化することができる。プロセス１１５０で、ｃｃフィールドの内容を、少なくとも１つの特徴として使用するためにオプションで抽出および／または正規化することができる。プロセス１１６０で、非無料電話番号を、オプションでメッセージの本文から抽出し、特徴として割り当てることもできる。非無料電話番号は、その電話番号のエリアコードおよび／または最初の３桁を使用してスパム送信者の位置をマッピングすることができるため、スパム送信者を識別するのに有用となる。複数の非無料電話番号がメッセージ内に存在する場合、プロセス１１６０で、各番号を別の特徴として抽出および使用することができる。

【0082】

同様に、オプションで、１つまたは複数のＵＲＬおよび／またはｍａｉｌｔｏ：リンク、あるいはその一部分を、それぞれプロセス１１７０およびプロセス１１８０にて抽出および／または正規化することができる。具体的には、ＵＲＬをパスウェイ分解（たとえば、ＵＲＬのファイル名部分）にかけることができ、ＵＲＬのＦＱＤＮ部分の末端に付けられている１つまたは複数の接尾語を分解することができる。これにより、パスウェイ内の接尾語の数に応じて、１つまたは複数の部分ＵＲＬが得られる。各部分ＵＲＬは、本発明による別の特徴として使用することができる。

【0083】

方法１１００は、引き続きメッセージの本体を走査し、本物のメッセージ内よりスパムメッセージ内で見つかる可能性の高い、またその逆である他の電子メールアドレス、ならびに（たとえば、先に選択されている、または決定されている）キーワードおよび／または句を捜すことができる。各単語または句は、機械学習システム、またはリストの要素、あるいはその両方のための特徴として抽出および使用することができる。

【0084】

先に説明しているように、インターネットを介して送信されるメッセージは、必要なサーバがわずか２つでもサーバからサーバに送信される可能性がある。メッセージを処理するサーバの数は、ファイアウォールと関連ネットワークアーキテクチャが存在する結果増大する。メッセージがサーバからサーバに渡されると、各サーバは、そのＩＰアドレスをｒｅｃｅｉｖｅｄ－ｆｒｏｍフィールドに付加する。また、各サーバは、先に付加されたどの発信元アドレスも修正する機能を有している。残念ながら、スパム送信者はこの機能を利用することができ、その位置および／または正体を偽装するため、およびメッセージの送信源について受信者を欺くため、偽物のアドレスをｒｅｃｅｉｖｅｄ－ｆｒｏｍフィールドに入力することができる。

【0085】

図１２は、着信メッセージのｒｅｃｅｉｖｅｄ－ｆｒｏｍ行内で本物と偽物（たとえば、スパム送信者）の付加サーバＩＰアドレスを区別するための例示的なプロセス１２００のフローチャートを示す。付加されたｒｅｃｅｉｖｅｄ－ｆｒｏｍアドレスは、追加された順序で調べることができる（たとえば、最初のものが最も最近追加されたものである）。したがって、プロセス１２１０で、ユーザは、一連の送信側サーバＩＰアドレスを遡って追跡し、最後の信頼されるサーバＩＰアドレスを決定することができる。プロセス１２２０で、最後の信頼されるサーバＩＰアドレス（システムのすぐ外側のもの）を、機械学習システムによって使用される特徴として抽出することができる。最後の信頼されるもの以後の他のＩＰアドレスは、どれも疑わしい、または信頼できないものと見なすことができ、無視することができるが、（概ね）問題のないＩＰアドレスのリストおよび（概ね）悪質なＩＰアドレスのリストに比較することも可能である。

【0086】

プロセス１２３０で、主張されている送信者のＦＱＤＮもまた、その送信者が本物であるか、それともスパム送信者であるか判定することを容易にするために抽出することができる。より具体的には、主張されているＦＱＤＮをドメイン分解によって分析し、複数の部分ＦＱＤＮに細分化することができる。たとえば、主張されているＦＱＤＮがａ．ｂ．ｃ．ｘ．ｃｏｍであると想像する。この主張されているＦＱＤＮは、ｂ．ｃ．ｘ．ｃｏｍ

→ｃ．ｘ．ｃｏｍ→ｘ．ｃｏｍ→ｃｏｍを生成するような形で分解されるはずである。したがって、各部分ＦＱＤＮセグメントならびに完全ＦＱＤＮを別の特徴として使用し、偽物の送信者と本物の送信者を決定する際に助けとすることができる。

【0087】

本発明はまた、ペアレンタル制御システムを使用することができる。ペアレンタル制御システムは、少なくとも部分的には、メッセージの一部の内容に基づいて、見るのに不適切なものとしてメッセージを分類し、不適切分類の理由を提供することができる。たとえば、（テキストまたは画像をベースとする）クリック可能なリンクとして、またはメッセージの本文内のテキストとして、ＵＲＬがメッセージ内に埋め込まれている可能性がある。ペアレンタル制御システムは、埋め込まれたＵＲＬと、その記憶された問題なし、および／または悪質ＵＲＬリストの１つまたは複数とを比較し、あるいはペアレンタル制御分類のための他の技法を使用して、メッセージの適切な分類を決定することができる。次いで、この分類は、機械学習システム内で、または特徴リスト上で、あるいはその両方で追加の特徴として使用することができる。

【0088】

図１３では、ペアレンタル制御システムの少なくとも１つの態様を本発明に組み込むための例示的なプロセス１３００のフローチャートが示されている。プロセス１３１０で１組のメッセージを受信した後に、プロセス１３２０で、ＵＲＬ、ｍａｉｌｔｏ：リンクもしくはｍａｉｌｔｏ：リンクに類似の他のテキスト、またはＵＲＬの一部分を求めてそのメッセージを走査することができる。プロセス１３３０で、メッセージが上記のどれかを含んでいるようには考えられない場合、プロセス１３００は、プロセス１３１０に戻る。しかし、メッセージがそのように考えられる場合、プロセス１３４０で、検出された特徴の少なくとも一部分を少なくとも１つのペアレンタル制御システムに渡すことができる。

【0089】

プロセス１３５０で、ペアレンタル制御システムは、ＵＲＬ、ｍａｉｌｔｏ：リンク、ＵＲＬサービス名、ＵＲＬパス、ＦＱＤＮ（たとえば、ＵＲＬ、電子メールアドレスなどのＦＱＤＮ部分など）の１つまたは複数のデータベースを調べることにより、ｍａｉｌｔｏ：リンク、ＵＲＬ、またはそれらの一部分を分類することができる。たとえば、メッセージを、ポルノ、「借金返済」、賭博、および他の同様な題材の少なくとも１つを含むものとして分類することができる。そのような分類は、プロセス１３６０で追加の特徴として抽出することができる。スパムメッセージの大部分の内容はそのような題材を含むため、ペアレンタル制御システムを組み込むことは、機械学習システムがトレーニングを行い、改良されたフィルタを構築するためにさらに使用することができる特徴を得る上で有用となる。他の分類も考えられるが、憎悪発言、性的題材、銃暴力、および麻薬関連題材も含まれ、そのような分類もまた、特徴として使用することができるものの、これには限られない。スパムメッセージは、これらの種類の題材に関連する内容を含むこともあり、含まないこともあるが、ユーザは、依然としてこれらの種類のメッセージを遮断したいと望む可能性がある。

【0090】

実際には、様々な分類により、様々なスパム度を示すことができる。たとえば、憎悪発言として分類されたメッセージは、（たとえば、スパムでない可能性が最も高いため）実質的にスパム度を示さない可能性がある。逆に、性的内容／題材として分類されたメッセージは、比較的高いスパム度を反映している可能性がある（たとえば、メッセージがスパムである９０％以下の確実性）。機械学習システムは、スパム度を反映するフィルタを構築することができる。したがって、ユーザ嗜好を満たすようにフィルタをカスタマイズおよび個別化することができる。

【0091】

すでに説明したように、無数の特徴をメッセージから抽出し、機械学習システムによるトレーニング用データとして、または、問題のない特徴および悪質な特徴を識別するリスト上の要素として使用することができる。特徴自体に加えて、特徴の質は、スパムを検出

および防止する上で有用となる。たとえば、ある特徴が送信者の電子メールアドレスであると考える。その電子メールアドレスをある特徴として使用することができ、新しい着信メッセージ内に現れるその電子メールアドレスの頻度または回数を別の特徴として使用することができるであろう。

【0092】

図１４は、この型の（たとえば、抽出された特徴の共通性または希少性に関する）特徴を抽出するための例示的なプロセス１４００のフローチャートを示す。スパム送信者は、その位置を急いで変更しようと頻繁に試み、その結果、たとえば以前見られなかったアドレスから、または以前は知られていなかった装置を指すＵＲＬを有するメールを送信する可能性がほとんどのユーザより高い。したがって、抽出された各特徴の型（たとえば、発信元ＩＰアドレス、ＵＲＬ、電子メールアドレス、ドメイン名など）について、各型の特徴のリストが保持されているものと仮定して、特定の特徴の発生率、頻度、または回数を追跡することができる。

【0093】

プロセス１４００は、プロセス１４１０で、着信メッセージからの１つまたは複数の特徴の抽出、および／または特徴の正規化から開始する。次いで、プロセス１４２０でその特徴を、先に抽出された、または複数の以前のメッセージ内で観察された特徴の１つまたは複数のリストに比較することができる。次いで、プロセス１４００は、現在の特徴が共通であるかどうか判定することができる。特徴の共通性は、最近の、および／または以前の着信メッセージ内にその特徴が現れる頻度を計算することによって決定することができる。プロセス１４３０で、そのメッセージが共通でない、またはそれほど共通でない（たとえば、共通性閾値を満たしていない）場合には、プロセス１４４０で、その希少性を追加の特徴として使用することができる。そうでない場合には、プロセス１４６０で、その特徴の共通性もまた特徴として使用することができる。

【0094】

上述した本発明によれば、以下の擬似コードを使用し、本発明の少なくとも１つの態様を実施することができる。変数名は、すべて大文字で示されている。追加の注意として、ａｄｄ－ｍａｃｈｉｎｅ－ｆｅａｔｕｒｅｓとａｄｄ－ｉｐ－ｆｅａｔｕｒｅｓという２つの関数が擬似コードの末尾で定義されている。「ＰＲＥＦＩＸ－ｍａｃｈｉｎｅ－ＭＡＣＨＩＮＥ」のような表記を使用し、ＰＲＥＦＩＸ変数内にあるものが「ｍａｃｈｉｎｅ」という単語に連結され、さらに「ｍａｃｈｉｎｅ」という単語がＭＡＣＨＩＮＥ変数内にあるものに連結されて構成されたストリングを示す。最後に、関数ａｄｄ－ｔｏ－ｆｅａｔｕｒｅ－ｌｉｓｔは、現在のメッセージに関連付けられた特徴のリストに特徴を書き込む。

【0095】

例示的な擬似コードは、次の通りである。

【0096】

【表5】

```
# for a given message, extract all the features

IPADDRESS := the last external IP address in the received-
from list;
```

【0097】

【表6】

```
add-ipfeatures(received, IPADDRESS);
SENDERS-ALLEGED-FQDN := FQDN in the last external IP
address in the received-from list;
add-machine-features(sendersfqdn, SENDERS-ALLEGED-FQDN);

for each email address type TYPE in (from, CC, to, reply-
to, embedded-mailto-link, embedded-address, and SMTP MAIL
FROM)
{
    for each address ADDRESS of type TYPE in the message {
        deobfuscate ADDRESS if necessary;
        add-to-feature-list TYPE-ADDRESS;
        if ADDRESS is of the form NAME@MACHINE then
        {
            add-machine-features(TYPE, MACHINE);
        }
        else
        { # ADDRESS is of form NAME@IPADDRESS
            add-ip-features(TYPE, IPADDRESS);
        }
    }
}

for each url type TYPE in (clickable-links, text-based-
links, embedded-image-links)
{
    for each URL in the message of type TYPE
    {
        deobfuscate URL;
        add-to-feature-list TYPE-URL;
        set PARENTALCLASS := parental control system class
of URL;
        add-to-feature-list TYPE-class-PARENTCLASS;
        while URL has a location suffix
        {
            remove location suffix from URL, i.e. x.y/a/b/c
-> x.y/a/b; x.y/a/b -> x.y/a; x.y/a;
        }
        # All suffixes have been removed; URL is now either
machine name or IP address
        if URL is machine name
        {
            add-machine-features(TYPE, URL);
        }
        else
        {
            add-ip-features(TYPE, URL);
```

【0098】

【表7】

```
        }
    }
}

function add-machine-features(PREFIX, MACHINE)
{
    add-ip-features(PREFIX-ip, nslookup(MACHINE));
    while MACHINE not equal ""
    {
        add-to-feature-list PREFIX-machine-MACHINE;
        remove beginning from MACHINE # (i.e. a.x.com ->
x.com, or x.com -> com);
    }
}

function add-ip-features(PREFIX, IPADDRESS)
{
    add-to-feature-list PREFIX-ipaddress-IPADDRESS;
    find netblock NETBLOCK of IPADDRESS;
    add-to-feature-list PREFIX-netblock-NETBLOCK;
    for N = 1 to 31 {
        MASKED = first N bits of IPADDRESS;
        add-to-feature-list PREFIX-masked-N-MASKED;
    }
}
```

【0099】
本発明の様々な態様についてさらに情報を提供するため、図１５および以下の考察では、本発明の様々な態様を実施することができる好適な動作環境１５１０を簡単に、一般的に説明するものとする。本発明について、１つまたは複数のコンピュータまたは他のデバイスによって実行されるプログラムモジュールなどのコンピュータ実行可能命令の一般的な文脈で説明されるが、本発明はまた、他のプログラムモジュールとの組合せで、および／またはハードウェアとソフトウェアの組合せとして実施することができることを、当業者であれば理解することができるであろう。

【0100】
ただし、一般に、プログラムモジュールは、特定のタスクを実行する、あるいは特定の抽象データ型を実施するルーチン、プログラム、オブジェクト、コンポーネント、データ構造などを含む。動作環境１５１０は、好適な動作環境の一例にすぎず、本発明の使用または機能の範囲についてどんな制限も暗示しないものとする。本発明と共に使用するのに適している可能性のある他の周知のコンピュータシステム、環境、および／または構成には、それだけには限らないが、パーソナルコンピュータ、ハンドヘルドデバイスまたはラップトップデバイス、マルチプロセッサシステム、マイクロプロセッサをベースとするシステム、プログラム可能な家電、ネットワークＰＣ、ミニコンピュータ、メインフレームコンピュータ、上記のシステムまたはデバイスを含む分散コンピューティング環境などが含まれる。

【0101】
図１５を参照すると、本発明の様々な態様を実施するための例示的な環境１５１０は、コンピュータ１５１２を含んでいる。コンピュータ１５１２は、処理装置１５１４、シス

テムメモリ１５１６、システムバス１５１８を含む。システムバス１５１８は、それだけには限らないが、システムメモリ１５１６を含むシステム構成要素を処理装置１５１４に結合する。処理装置１５１４は、様々な使用可能なプロセッサのいずれかとすることができる。デュアルマイクロプロセッサおよび他のマルチプロセッサアーキテクチャもまた、処理装置１５１４として使用することができる。

【0102】

システムバス１５１８は、メモリバスもしくはメモリコントローラ、周辺機器バスもしくは外部バス、および／または任意の様々な使用可能なバスアーキテクチャを使用するローカルバスを含めて、いくつかの型のバス構造のいずれかとすることができる。バスアーキテクチャには、それだけには限らないが、１１ビットバス、ＩＳＡ（Ｉｎｄｕｓｔｒｉａｌ　Ｓｔａｎｄａｒｄ　Ａｒｃｈｉｔｅｃｔｕｒｅ）バス、ＭＣＡ（Ｍｉｃｒｏ－Ｃｈａｎｎｅｌ　Ａｒｃｈｉｔｅｃｔｕｒｅ）バス、ＥＩＳＡ（Ｅｘｔｅｎｄｅｄ－ＩＳＡ）バス、ＩＤＥ（Ｉｎｔｅｌｌｉｇｅｎｔ　Ｄｒｉｖｅ　Ｅｌｅｃｔｒｏｎｉｃｃｓ）、ＶＥＳＡローカルバス（ＶＬＢ）、ＰＣＩ（Ｐｅｒｉｐｈｅｒａｌ　Ｃｏｍｐｏｎｅｎｔ　Ｉｎｔｅｒｃｏｎｎｅｃｔ）、ユニバーサルシリアルバス（ＵＳＢ）、ＡＧＰ（Ａｄｖａｎｃｅｄ　Ｇｒａｐｈｉｃｓ　Ｐｏｒｔ）、ＰＣＭＣＩＡ（Ｐｅｒｓｏｎａｌ　Ｃｏｍｐｕｔｅｒ　Ｍｅｍｏｒｙ　Ｃａｒｄ　Ｉｎｔｅｒｎａｔｉｏｎａｌ　Ａｓｓｏｃｉａｔｉｏｎ）バス、ＳＣＳＩ（Ｓｍａｌｌ　Ｃｏｍｐｕｔｅｒ　Ｓｙｓｔｅｍｓ　Ｉｎｔｅｒｆａｃｅ）が含まれる。

【0103】

システムメモリ１５１６には、揮発性メモリ１５２０および不揮発性メモリ１５２２が含まれる。起動中などにコンピュータ１５１２内の要素間で情報を転送する基本ルーチンを含む基本入出力システム（ＢＩＯＳ）は、不揮発性メモリ１５２２内に記憶されている。限定ではなく例を挙げると、不揮発性メモリ１５２２には、読出し専用メモリ（ＲＯＭ）、プログラム可能なＲＯＭ（ＰＲＯＭ）、電気的プログラム可能なＲＯＭ（ＥＰＲＯＭ）、電気的消去可能なＲＯＭ（ＥＥＰＲＯＭ）、またはフラッシュメモリが含まれる。揮発性メモリ１５２０には、外部キャッシュメモリとして動作するランダムアクセスメモリ（ＲＡＭ）が含まれる。限定ではなく例を挙げると、ＲＡＭは、スタティックＲＡＭ（ＳＲＡＭ）、ダイナミックＲＡＭ（ＤＲＡＭ）、シンクロナスＤＲＡＭ（ＳＤＲＡＭ）、ダブルデータレートＳＤＲＡＭ（ＤＤＲ　ＳＤＲＡＭ）、ＥＳＤＲＡＭ（ｅｎｈａｎｃｅｄ　ＳＤＲＡＭ）、ＳＬＤＲＡＭ（Ｓｙｎｃｈｌｉｎｋ　ＤＲＡＭ）、ダイレクトラムバスＲＡＭ（ＤＲＲＡＭ）など、多数の形態で使用可能である。

【0104】

コンピュータ１５１２はまた、取外し可能／固定、揮発性／不揮発性コンピュータ記憶媒体を含む。図１５は、たとえば、ディスク記憶装置１５２４を示す。ディスク記憶装置１５２４には、それだけには限らないが、磁気ディスクドライブ、フロッピー（登録商標）ディスクドライブ、テープドライブ、Ｊａｚドライブ、Ｚｉｐドライブ、ＬＳ－１００ドライブ、フラッシュメモリカード、メモリースティックのようなデバイスが含まれる。さらに、ディスク記憶装置１５２４には、それだけには限らないが、コンパクトディスクＲＯＭデバイス（ＣＤ－ＲＯＭ）、記録可能なＣＤドライブ（ＣＤ－Ｒドライブ）、再書込み可能なＣＤドライブ（ＣＤ－ＲＷドライブ）、またはデジタル多用途ディスクＲＯＭドライブ（ＤＶＤ－ＲＯＭ）など光ディスクドライブを含めて、記憶媒体が別個に、または他の記憶媒体との組合せで含まれる可能性がある。ディスク記憶装置１５２４のシステムバス１５１８に対する接続を容易にするために、インターフェース１５２６など取外し可能または固定インターフェースが一般に使用される。

【0105】

図１５は、ユーザと、好適な動作環境１５１０に説明される基本的なコンピュータ資源との間を媒介するものとして動作するソフトウェアについて述べていることを理解されたい。そのようなソフトウェアには、オペレーティングシステム１５２８が含まれる。オペレーティングシステム１５２８は、ディスク記憶装置１５２４に記憶することができ、コ

ンピュータシステム１５１２の資源を制御し、割り振るように動作する。システムアプリケーション１５３０は、システムメモリ１５１６内またはディスク記憶装置１５２４に記憶されたプログラムモジュール１５３２およびプログラムデータ１５３４を介して、オペレーティングシステム１５２８による資源の管理を利用する。本発明が様々なオペレーティングシステムまたはオペレーティングシステムの組合せと共に実施することができることを理解されたい。

【0106】

ユーザは、入力デバイス１５３６を介してコンピュータ１５１２にコマンドまたは情報を入力する。入力デバイス１５３６には、それだけには限らないが、マウスなどポインティングデバイス、トラックボール、スタイラス、タッチパッド、キーボード、マイクロフォン、ジョイスティック、ゲームパッド、衛星パラボラアンテナ、スキャナ、ＴＶ同調器カード、デジタルカメラ、デジタルビデオカメラ、ウェブカメラなどが含まれる。これら、および他の入力デバイスは、インターフェースポート１５３８を介して、システムバス１５１８を通じて処理装置１５１４に接続する。インターフェースポート１５３８には、たとえば、シリアルポート、パラレルポート、ゲームポート、ユニバーサルシリアルバス（ＵＳＢ）が含まれる。出力デバイス１５４０は、入力デバイス１５３６と同じ型のポートのいくつかを使用する。したがって、たとえばＵＳＢポートは、コンピュータ１５１２に入力を提供するために、また、コンピュータ１５１２から出力デバイス１５４０に情報を出力するために使用することができる。出力アダプタ１５４２は、出力デバイス１５４０の中でも、特別なアダプタを必要とするモニタ、スピーカ、プリンタのようないくつかの出力デバイス１５４０があることを示すために提供されている。限定ではなく例を挙げると、出力アダプタ１５４２には、出力デバイス１５４０とシステムバス１５１８の間で出力手段を提供するビデオカードおよびサウンドカードが含まれる。他のデバイスおよび／またはデバイスのシステムは、リモートコンピュータ１５４４など、入力機能と出力機能を共に提供することに留意されたい。

【0107】

コンピュータ１５１２は、リモートコンピュータ１５４４など、１つまたは複数のリモートコンピュータに対する論理接続を使用してネットワーク環境内で動作することができる。リモートコンピュータ１５４４は、パーソナルコンピュータ、サーバ、ルータ、ネットワークＰＣ、ワークステーション、マイクロプロセッサをベースとする家電、ピアデバイスまたは他の通常のネットワークノードなどとすることができ、一般に、コンピュータ１５１２に関して上述した要素の多数または全部を含む。簡潔にするために、メモリ記憶装置１５４６だけがリモートコンピュータ１５４４と共に示されている。リモートコンピュータ１５４４は、ネットワークインターフェース１５４８を介してコンピュータ１５１２に論理的に接続され、次いで、通信接続１５５０を介して物理的に接続される。ネットワークインターフェース１５４８は、ローカルエリアネットワーク（ＬＡＮ）および広域ネットワーク（ＷＡＮ）など、通信ネットワークを含む。ＬＡＮ技術には、光ファイバ分散データインターフェース（ＦＤＤＩ）、銅分散データインターフェース（ＣＤＤＩ）、イーサネット（登録商標）／ＩＥＥＥ１１０２．３、トークンリング／ＩＥＥＥ１１０２．５などが含まれる。ＷＡＮ技術には、それだけには限らないが、ポイント・トゥ・ポイント・リンク、総合デジタル通信網（ＩＳＤＮ）のような回路交換ネットワークとその変形形態、パケット交換ネットワーク、およびデジタル加入者回線（ＤＳＬ）が含まれる。

【0108】

通信接続１５５０は、ネットワークインターフェース１５４８をバス１５１８に接続するために使用されるハードウェア／ソフトウェアを指す。通信接続１５５０は、図が見やすいようにコンピュータ１５１２の内側で示されているが、コンピュータ１５１２の外側とすることもできる。例示する目的にすぎないが、ネットワークインターフェース１５４８に対する接続に必要なハードウェア／ソフトウェアには、通常の電話級モデム、ケーブルモデムおよびＤＳＬモデムを含むモデム、ＩＳＤＮアダプタ、ならびにイーサネット（登録商標）カードなど、内部技術および外部技術が含まれる。

【0109】
上述したものには、本発明の諸例が含まれる。当然ながら、本発明について説明するために構成要素または方法の考えられるあらゆる組合せについて説明することは可能でなく、本発明の多数の他の組合せおよび変形が可能であることを、当業者なら理解することができる。したがって、本発明は、添付の特許請求の範囲の精神および範囲内に入るそのような変更、修正、および変形形態をすべて包含するものとする。さらに「ｉｎｃｌｕｄｅｓ（含む）」という用語が詳細な説明または特許請求の範囲で使用されている限り、そのような用語は、「ｃｏｍｐｒｉｓｉｎｇ（含む、備える）」が特許請求の範囲内で転換句として使用されたとき解釈されるように「ｃｏｍｐｒｉｓｉｎｇ」という用語と同様に包括的であるものとする。

【図面の簡単な説明】

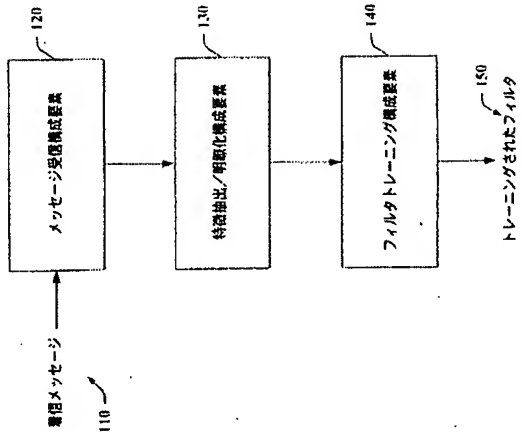
【0110】
【図1】本発明の態様による、スパム防止を容易にするシステムの高レベルブロック図である。
【図2】本発明の態様による、着信メッセージから１つまたは複数の特徴を抽出することによってスパム防止を容易にするシステムのブロック図である。
【図3】本発明の態様による、ＩＰアドレスから抽出することができる複数の特徴の概略図である。
【図4】本発明の態様による、ＦＱＤＮから抽出することができる複数の特徴の概略図である。
【図5】本発明の態様による、電子メールアドレスから抽出することができる複数の特徴の概略図である。
【図6】本発明の態様による、ＵＲＬまたはウェブアドレスから抽出することができる複数の特徴の概略図である。
【図7】本発明の態様による、フィルタをトレーニングすることに関連する例示的な方法のフローチャートである。
【図8】本発明の態様による、トレーニングされたフィルタを使用することに関連する例示的な方法のフローチャートである。
【図9】本発明の態様による、リストを作成することに関連する例示的な方法のフローチャートである。
【図10】本発明の態様による、フィルタをトレーニングするためにリストを使用することに関連する例示的な方法のフローチャートである。
【図11】本発明の態様による、少なくとも図７および図８の方法の中で参照された工程のフローチャートである。
【図12】本発明の態様による、本物と偽物の発信元（ｒｅｃｅｉｖｅｄ－ｆｒｏｍ）ＩＰアドレスを区別することを容易にする工程のフローチャートである。
【図13】本発明の態様による、着信メッセージから特徴を生成および／または抽出する際にペアレンタル制御システムを組み込む方法のフローチャートである。
【図14】本発明の態様による、機械学習システム内で使用される特徴セットの作成を容易にする方法のフローチャートである。
【図15】本発明の様々な態様を実施するための例示的な環境の図である。

【符号の説明】

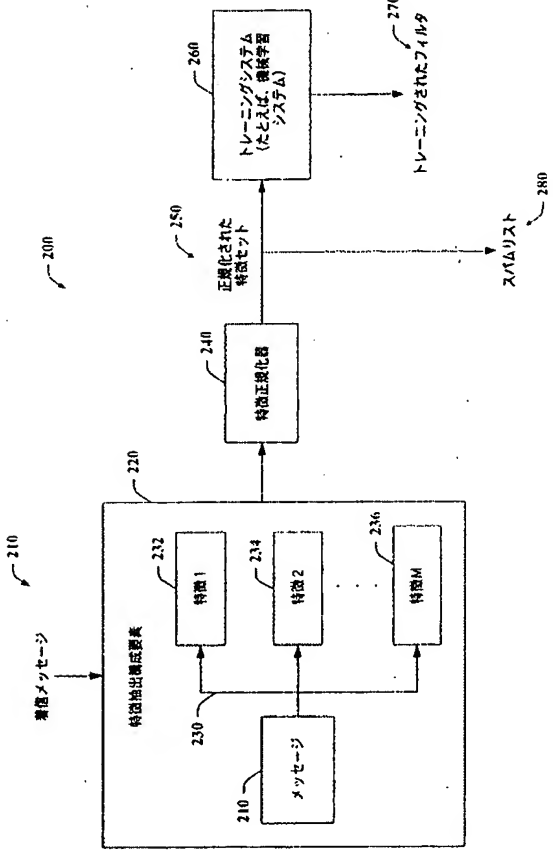
【0111】
１１０　着信メッセージ
１２０　メッセージ受信構成要素
１３０　特徴抽出／不明瞭構成要素
１４０　フィルタトレーニング構成要素
１５０　トレーニングされたフィルタ
２１０　着信メッセージ、メッセージ
２２０　特徴抽出構成要素

２３０　特徴
２３２　特徴１
２３４　特徴２
２３６　特徴Ｍ
２４０　特徴正規化器
２５０　正規化された特徴セット
２６０　トレーニングシステム
２７０　トレーニングされたフィルタ
２８０　スパムリスト
３００　ＩＰアドレス
３０２　ブロックＩＤ
３１４　ＩＰアドレスの階層
３１６　特徴の希少性
４００　ＦＱＤＮ
４０２　ホスト名
４０４　ドメイン名
４０６　内容
４１４　特徴の型
５００　電子メールアドレス
５０２　ユーザ名
５０４　ＦＱＤＮ
６００　ＵＲＬ
１５１０　環境
１５１２　コンピュータ
１５１４　処理装置
１５１６　システムメモリ
１５１８　システムバス
１５２０　揮発性メモリ
１５２２　不揮発性メモリ
１５２４　ディスク記憶装置
１５２６　インターフェース
１５２８　オペレーティングシステム
１５３０　アプリケーション
１５３２　モジュール
１５３４　データ
１５３６　入力デバイス
１５３８　インターフェースポート
１５４０　出力デバイス
１５４２　出力アダプタ
１５４４　リモートコンピュータ
１５４６　メモリ記憶装置
１５４８　ネットワークインターフェース
１５５０　通信接続

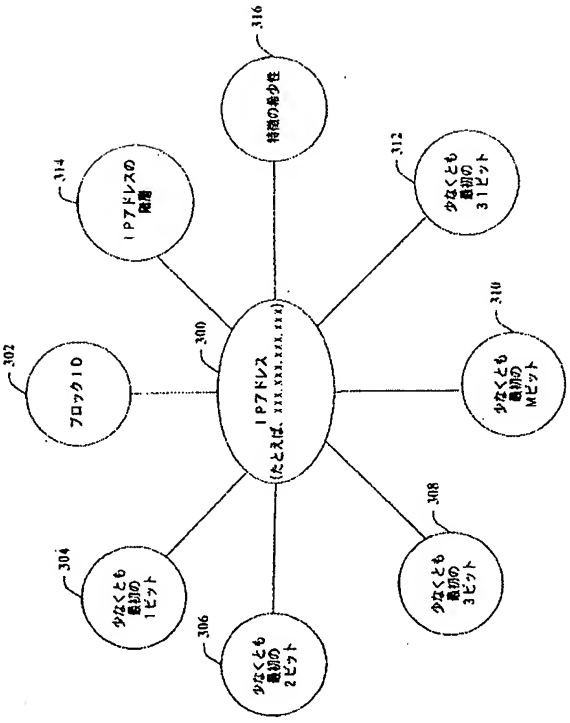
【図1】


メッセージ受信構成要素
特徴抽出／特徴化構成要素
フィルタトレーニング構成要素
トレーニングされたフィルタ
着信メッセージ
120
130
140
150
110

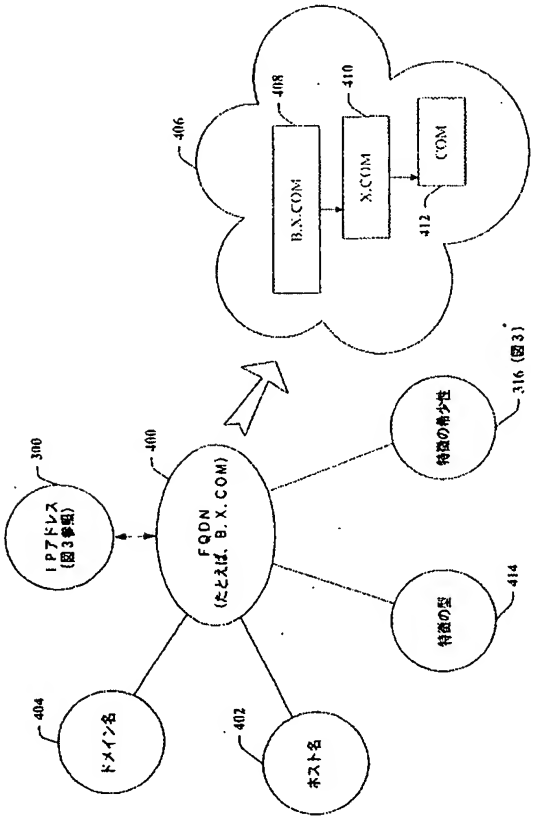
【図2】


トレーニングシステム（たとえば、機械学習システム）
トレーニングされたフィルタ
正規化された特徴セット
スパムリスト
特徴正規化器
特徴抽出構成要素
着信メッセージ
特徴1
特徴2
特徴M
メッセージ
200
210
220
230
232
234
236
240
250
260
270
280

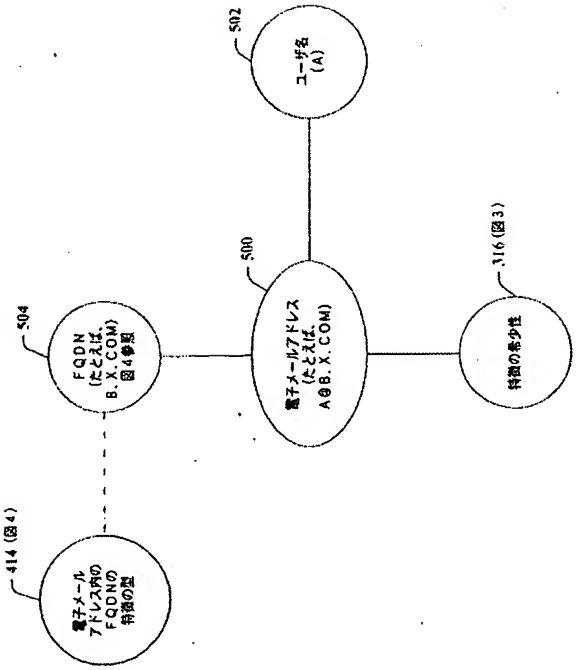
【図3】


IPアドレス（たとえば、xxx.xxx.xxx.xxx）
ブロックID
IPアドレスの階層
特徴の希少性
少なくとも最初の1ビット
少なくとも最初の2ビット
少なくとも最初の3ビット
少なくとも最初の31ビット
少なくとも最初のMビット
300
302
304
306
308
310
312
314
316

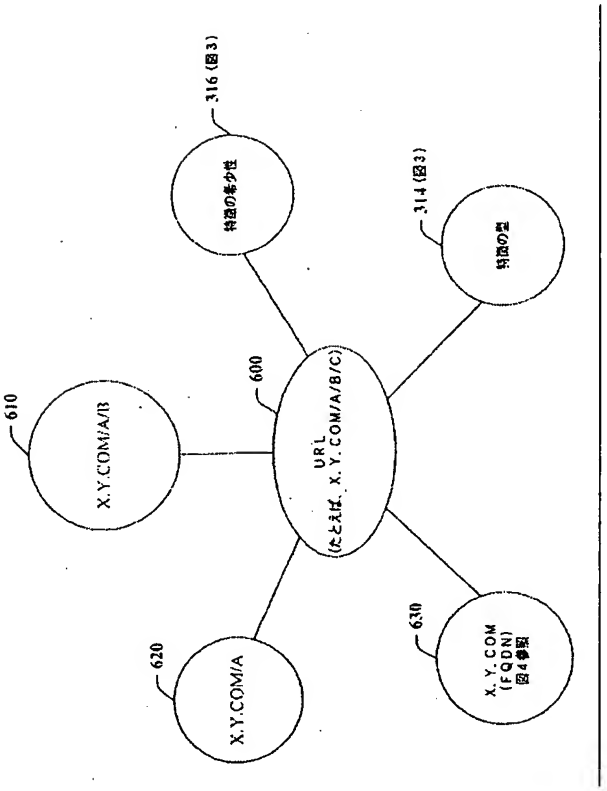
【図4】


B.X.COM
X.COM
COM
FQDN（たとえば、B.X.COM）
IPアドレス（図3参照）
ドメイン名
ホスト名
特徴の希少性
特徴の型
406
408
410
412
400
300
404
402
316（図3）
414

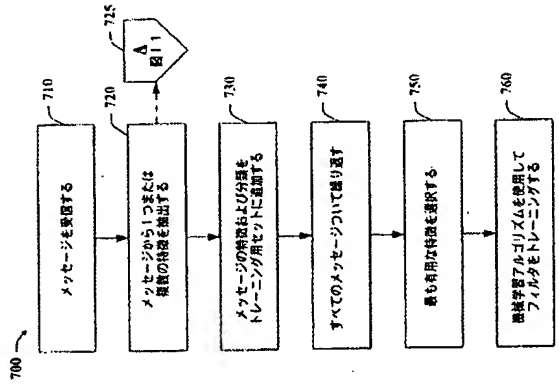
【図5】


502
ユーザ名
(A)
500
電子メールアドレス
(たとえば、
A@B.X.COM)
504
FQDN
(たとえば、
B.X.COM)
図4参照
316 (図3)
特徴の希少性
414 (図4)
電子メール
アドレス内の
FQDNの
特徴の型

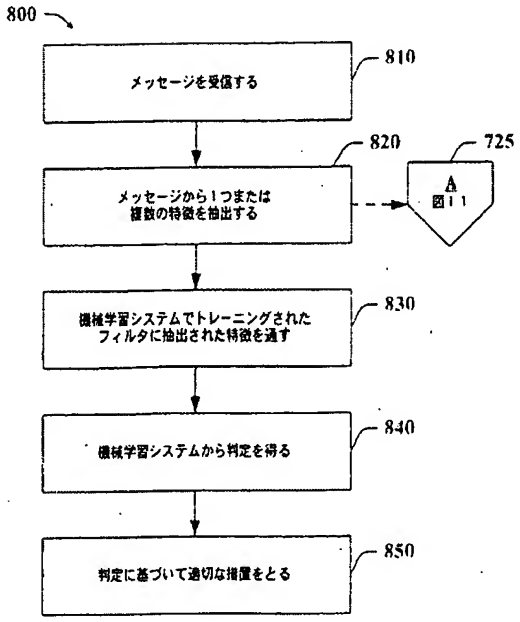
【図6】


316 (図3)
特徴の希少性
314 (図3)
特徴の型
600
URL
(たとえば、X.Y.COM/A/B/C)
610
X.Y.COM/A/B
620
X.Y.COM/A
630
X.Y.COM
(FQDN)
図4参照

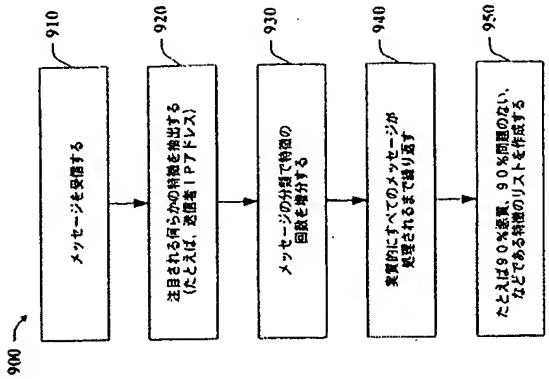
【図7】


700
710
メッセージを受信する
720
メッセージから1つまたは
複数の特徴を抽出する
725
A
図11
730
メッセージの特徴およびその値を
トレーニング用セットに追加する
740
すべてのメッセージについて繰り返す
750
最も有用な特徴を選択する
760
機械学習アルゴリズムを使用して
フィルタをトレーニングする

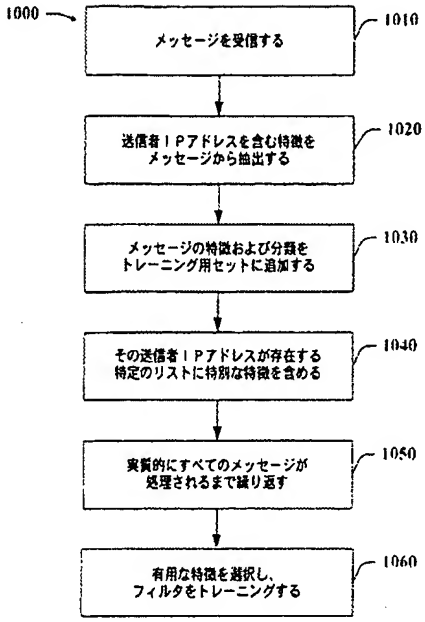
【図8】


800
メッセージを受信する
810
820
メッセージから1つまたは
複数の特徴を抽出する
725
A
図11
機械学習システムでトレーニングされた
フィルタに抽出された特徴を通す
830
機械学習システムから判定を得る
840
判定に基づいて適切な措置をとる
850

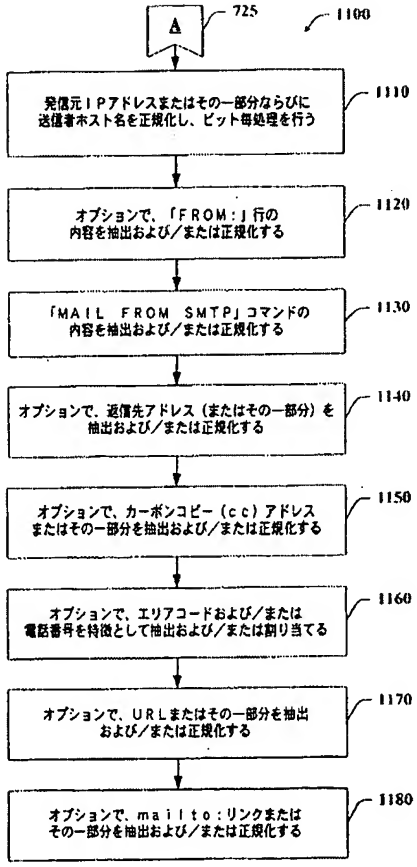
【図9】



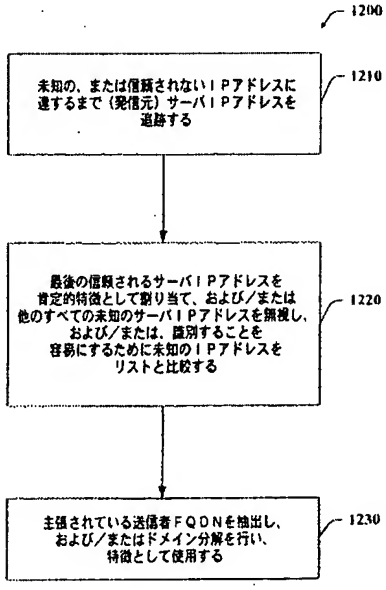
【図10】



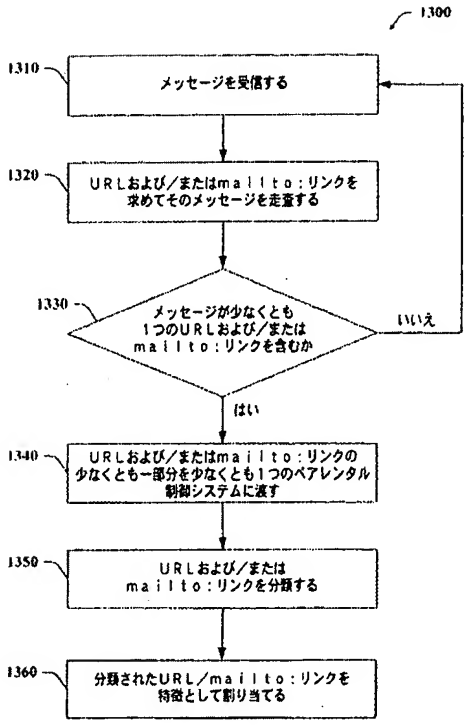
【図11】



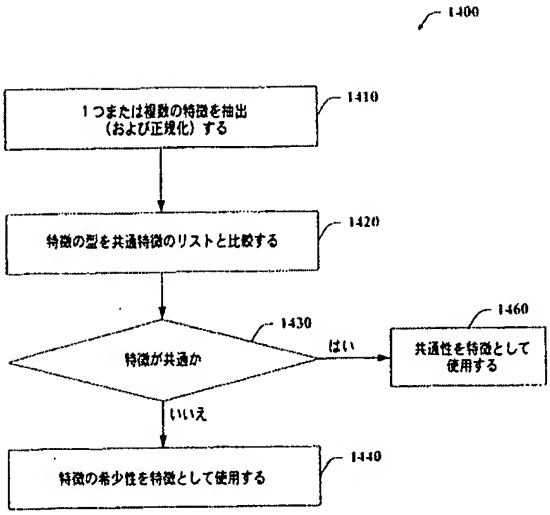
【図12】



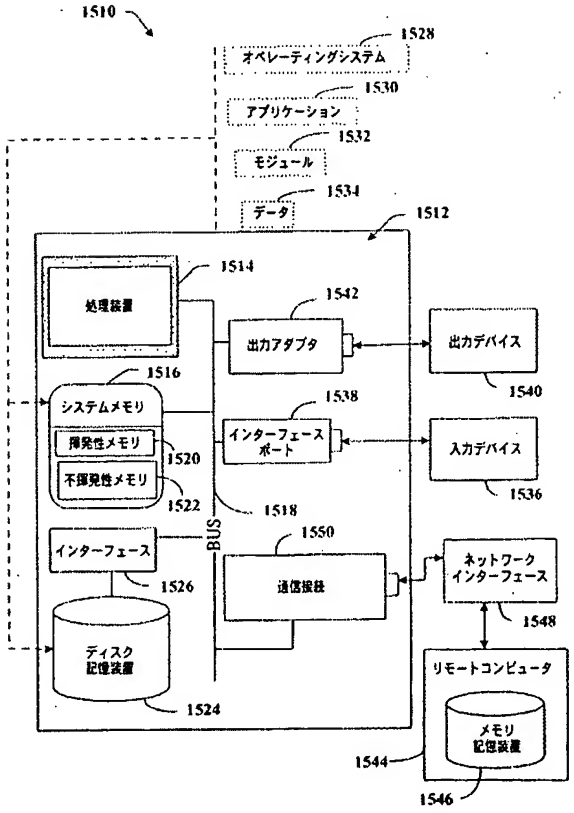
【図13】


1300
1310
メッセージを受信する
1320
URLおよび/またはmailto:リンクを求めてそのメッセージを走査する
1330
メッセージが少なくとも1つのURLおよび/またはmailto:リンクを含むか
いいえ
はい
1340
URLおよび/またはmailto:リンクの少なくとも一部分を少なくとも1つのペアレンタル制御システムに渡す
1350
URLおよび/またはmailto:リンクを分類する
1360
分類されたURL/mailto:リンクを特徴として割り当てる

【図14】


1400
1つまたは複数の特徴を抽出（および正規化）する
1410
特徴の型を共通特徴のリストと比較する
1420
特徴が共通か
1430
はい
共通性を特徴として使用する
1460
いいえ
特徴の希少性を特徴として使用する
1440

【図15】


1510
オペレーティングシステム
1528
アプリケーション
1530
モジュール
1532
データ
1534
1512
処理装置
1514
出力アダプタ
1542
出力デバイス
1540
システムメモリ
1516
揮発性メモリ
1520
不揮発性メモリ
1522
インターフェースポート
1538
入力デバイス
1536
1518
BUS
インターフェース
1526
通信接続
1550
ネットワークインターフェース
1548
ディスク記憶装置
1524
リモートコンピュータ
メモリ記憶装置
1544
1546

(72)発明者　ロバート　エル．ラウンスウェイト
アメリカ合衆国　９８０２４　ワシントン州　フォール　シティ　２８７　アベニュー　サウスイースト　４１４８
(72)発明者　ダニエル　グォゼズ
アメリカ合衆国　９８０７４　ワシントン州　サマミッシュ　ノースイースト　２３　コート　２０５３３
(72)発明者　ジョン　ディー．メヘル
アメリカ合衆国　９８１０３　ワシントン州　シアトル　ホイットマン　アベニュー　ノース　３６２４　ナンバー３
(72)発明者　ネイサン　ディー．ハウエル
アメリカ合衆国　９８１３３　ワシントン州　シアトル　ノース　１０５　ストリート　９３９　アパートメント　エー
(72)発明者　ミカ　シー．ルパーズバーグ
アメリカ合衆国　９８１２２　ワシントン州　シアトル　イースト　パイン　ストリート　４１７　ナンバー２０９
(72)発明者　ブライアン　ティー．スターバック
アメリカ合衆国　９８０１９　ワシントン州　ドボル　ノースイースト　１４０　コート　２７５１７

【要約の続き】

【外国語明細書】
2004362559000001.pdf